

---

# **Estimation of Continuous–Time Financial Models Using High–Frequency Data**

---

Dissertation an der Fakultät für Mathematik, Informatik und Statistik der  
Ludwig-Maximilians-Universität München

vorgelegt von

**Christian Pigorsch am 1. Februar 2007**

Ludwig-Maximilians-Universität München  
Fakultät für Mathematik, Informatik und Statistik

Dissertation

# **Estimation of Continuous–Time Financial Models Using High–Frequency Data**

vorgelegt von  
Christian Pigorsch

München, den 1. Februar 2007

Erstgutachter: Prof. Stefan Mittnik, Ph.D.  
Zweitgutachter: Prof. Dr. Ludwig Fahrmeir  
Externer Gutachter: Prof. A. Ronald Gallant, Ph.D.

Rigorosum: 5. Juni 2007

# Contents

<b>1</b>	<b>Introduction</b>	<b>8</b>
<b>2</b>	<b>High-Frequency Information</b>	<b>11</b>
2.1	Definition of Realized Variation and Covariation Measures . . . . .	12
2.1.1	Realized Variation . . . . .	14
2.1.2	Realized Covariation . . . . .	16
2.2	Stylized Facts of Returns and Realized Variation Measures . . . . .	17
2.2.1	Univariate Dataset . . . . .	17
2.2.2	Multivariate Dataset . . . . .	24
<b>3</b>	<b>Statistical Assessment of Univariate Continuous-Time Stochastic Volatility Models</b>	<b>36</b>
3.1	Model Specifications . . . . .	37
3.1.1	Affine Models . . . . .	38
3.1.2	Logarithmic Models . . . . .	39
3.1.3	Jump-Diffusion Models . . . . .	39
3.1.4	Model Definitions . . . . .	40
3.2	Estimation Methodology . . . . .	40
3.2.1	The General Scientific Modeling Method . . . . .	42
3.3	The Auxiliary Model . . . . .	46
3.3.1	A Discrete-Time Model for Daily Returns and Realized Variations . . . . .	47
3.3.2	Equation-by-Equation Estimation . . . . .	50
3.3.3	System Estimation . . . . .	62
3.3.4	Further Accuracy Checks via Simulations . . . . .	65
3.4	Prior Information . . . . .	73
3.5	Empirical Results . . . . .	74
3.6	Summary . . . . .	84
<b>4</b>	<b>A Multivariate Extension of the Ornstein-Uhlenbeck Stochastic Volatility Model</b>	<b>86</b>
4.1	The Univariate Non-Gaussian OU-Type Stochastic Volatility Model	88
4.2	Positive Semidefinite Processes of OU-Type . . . . .	90
4.2.1	Notation . . . . .	91
4.2.2	Definition and Probabilistic Properties . . . . .	92
4.2.3	The Integrated Process . . . . .	95

## Contents

4.2.4	Marginal Dynamics . . . . .	96
4.3	The Multivariate OU-Type Stochastic Volatility Model . . . . .	97
4.3.1	Second Order Structure . . . . .	98
4.3.2	State Space Representation . . . . .	103
4.3.3	Realized Quadratic Variation . . . . .	106
4.4	Estimation Methods and Finite Sample Properties . . . . .	108
4.4.1	Estimation Methods . . . . .	108
4.4.2	Monte-Carlo Analysis . . . . .	110
4.5	Empirical Application . . . . .	114
4.6	Summary . . . . .	123
<b>5</b>	<b>Conclusion</b>	<b>125</b>

# List of Tables

2.1	Descriptive Statistics of the Univariate Dataset . . . . .	21
2.2	Company Descriptions of the Multivariate Dataset . . . . .	26
2.3	Description of the Multivariate Dataset . . . . .	27
2.4	Descriptive Statistics of the Multivariate Dataset (C, INTC, MSFT, PFE) . . . . .	34
3.1	Jump–Diffusion Model Specifications . . . . .	41
3.2	Single–Equation Estimation Results of the Auxiliary Model . . . . .	51
3.3	System Estimation Results of the Auxiliary Model . . . . .	64
3.4	Restricted System Estimation Results of the Auxiliary Model . . . . .	66
3.5	Simulation Results . . . . .	70
3.6	Estimation Results of the Continuous–Time Stochastic Volatility Models . . . . .	75
3.7	Summary Statistics of Model–Implied Distributions . . . . .	79
3.8	Summary Statistics of Model–Implied Conditional Distributions . . . . .	82
4.1	Monte–Carlo Results . . . . .	111
4.2	Univariate Estimation Results for MSFT . . . . .	116
4.3	Univariate Estimation Results for INTC . . . . .	117
4.4	Bivariate Estimation Results for MSFT and INTC . . . . .	120
4.5	Bivariate Estimation Results for MSFT and INTC, Characteristics . . . . .	121

# List of Figures

2.1	Volatility–Signature Plot of the S&P500 Index Futures . . . . .	19
2.2	Time Series of Returns, Logarithmic Realized Variance, Logarithmic Bipower Variation and Jumps . . . . .	20
2.3	Unconditional Distributions of Standardized Returns, Logarithmic Realized Variance, Logarithmic Bipower Variation and Jumps . . . .	22
2.4	Sample Autocorrelations and Partial Autocorrelations of Returns, Logarithmic Realized Variance, Logarithmic Bipower Variation and Jumps . . . . .	23
2.5	News–Impact Curves for Logarithmic Realized Variance, Logarithmic Bipower Variation and Jumps . . . . .	25
2.6	U–shaped Intraday Patterns . . . . .	28
2.7	Autocovariance Function of the Raw and Adjusted Returns . . . . .	30
2.8	Daily Returns and Logarithmic Realized Variances . . . . .	32
2.9	Daily Realized Correlations . . . . .	33
3.1	Residual Analysis of the (log.) Bipower Variation Equation . . . . .	52
3.2	Residual Analysis of the Jump Equation . . . . .	53
3.3	Residual Analysis of the Return Equation . . . . .	54
3.4	The Volatility of Bipower Variation . . . . .	55
3.5	Dependency Analysis of the Residuals between the Return Equation and Bipower Variation Equation . . . . .	58
3.6	Dependency Analysis of the Residuals between the Return Equation and Jump Equation . . . . .	59
3.7	Dependency Analysis of the Residuals between the Bipower Variation Equation and the Jump Equation . . . . .	60
3.8	CDF Scatter Plot of the Single–Equation Innovations . . . . .	61
3.9	CDF Scatter Plot of the System Innovations . . . . .	67
3.10	Simulated Paths . . . . .	69
3.11	Sample Quantiles of Returns, Logarithmic Realized Variance, Logarithmic Bipower Variation and Jumps . . . . .	71
3.12	Sample Autocorrelations and Partial Autocorrelations of Returns, Logarithmic Realized Variance, Logarithmic Bipower Variation and Jumps . . . . .	72
3.13	Sample Autocorrelations and Partial Autocorrelations of Realized Variance and Bipower Variation both in Standard Deviation Form .	73

## *List of Figures*

3.14	Unconditional Distributions of the Mean of the Returns, Realized Variance and Bipower Variation . . . . .	77
3.15	Unconditional Distributions of the Mean of the Jump Measure, Correlation and the Ljung–Box Statistics . . . . .	78
4.1	Simulated Univariate Sample Path . . . . .	89
4.2	Simulated Bivariate Sample Path . . . . .	99
4.3	Simulated Bivariate Sample Path, Realized Correlation and Scatter Plot . . . . .	100
4.4	Simulated Distributions of the Parameter Estimates . . . . .	112
4.5	Simulated Distributions of Implied Daily Return Characteristics . .	113
4.6	Model–Implied and Empirical Daily Autocorrelation Functions for MSFT . . . . .	118
4.7	Model–Implied and Empirical Daily Autocorrelation Functions for INTC . . . . .	119
4.8	Model–Implied and Empirical Daily Autocorrelation Functions Based on the Bivariate Estimation Results for MSFT and INTC . . . . .	122

# 1 Introduction

Modeling the dynamics of asset prices and in particular financial volatility is crucial for derivative pricing, risk management applications, and asset allocation decisions. With the recent availability of high-frequency, or tick-by-tick transaction, data of various financial markets the research in this area has taken new avenues. In particular, the new information contained in the high-frequency returns is exploited for example for the direct modeling of these high-frequency returns, as well as for the construction and modeling of lower-frequency nonparametric volatility measures.

On the intraday level, high-frequency data revealed that returns are subject to market microstructure frictions, such as transaction costs or bid-and-ask spreads, and other specific intraday patterns such as the U-shaped volatility over the day, or lunch-time effects. The existence of such effects complicates the direct modeling of high-frequency returns and the literature therefore focuses on modeling the realized variation measures, which effectively summarize on a lower level the most important information inherent in the high-frequency data. In fact, this literature builds on the general result that under ideal conditions the sum over the outer product of successively more finely sampled high-frequency returns converges to the quadratic variation of the price process (see Andersen and Bollerslev, 1998; Andersen et al., 2001b; Barndorff-Nielsen and Shephard, 2002b), an idea that already dates back to Merton (1980). However, the recent theoretical developments also allow the decomposition of the quadratic variation into the variation coming from the continuous-sample-path evolution of the price process, as measured by the so-called Bipower variation first introduced by Barndorff-Nielsen and Shephard (2004b, 2005), and the variation coming from the jumps. As such, these measures provide new information on the distribution and dynamics of the two volatility components, as well as on the importance of jumps, which in turn can be useful for modeling the dynamics of the price and volatility processes. Chapter 2 of this thesis provides a detailed discussion on the definition and construction of the realized (co)variation measures and investigates their empirical properties, which are exploited in the subsequent chapters for the statistical assessment of continuous-time stochastic volatility models.

In the finance and econometrics literature the continuous-time stochastic volatility models play a major role for asset pricing and risk management. Furthermore, due to their continuous-time formulation these models are informative about the price process at any frequency. As a consequence a plethora of different continuous-time stochastic volatility models has been developed, including for example the affine and logarithmic jump-diffusion models (e.g. Andersen et al., 2002; Chernov et al., 2003; Eraker, 2001; Eraker et al., 2003) in which the volatility



and the price processes are driven by jump–diffusion processes; the non–Gaussian Ornstein–Uhlenbeck–type models of Barndorff-Nielsen and Shephard (2001b), in which the volatility is modeled by a pure jump process; the Lévy–driven continuous AR(FI)MA stochastic volatility models of e.g. Brockwell (2001) and Marquardt (2004), which allow for a more flexible structure in the autocorrelation function of the returns; as well as for example the time–changed Lévy processes (e.g. Carr et al., 2003; Huang and Wu, 2004), in which stochastic volatility is introduced by exchanging calendar time with economic time.

Given the large number of different types of continuous–time stochastic volatility models it is interesting to assess their ability to reproduce the stylized facts of stock returns, and to compare their implied empirical properties. However, the empirical validation of these models is complicated by the existence of unobserved state variables, the rare availability of the transition density, and the discreteness of the observed prices. Consequently, to overcome these problems different estimation strategies have been developed and applied, such as simulated maximum likelihoods methods, MCMC methods and indirect inference approaches. Most of the corresponding empirical studies are based on daily or lower–frequency data and the empirical results typically do not allow for a very clear distinction between the different models. Importantly, they do not allow the distinction between pure diffusion multi–factor stochastic volatility models and lower–order models with jumps. In view of the often observed large intraday price movements, however, one might conjecture that the daily data most frequently used in the estimation of these models may simply not be informative enough to provide a firm answer.

In this thesis we therefore use high–frequency financial data and re–assess the adequacy of the continuous–time stochastic volatility models. More specifically, as the direct estimation of specific parametric volatility models with large samples of high–frequency intraday data remains extremely challenging from a computational perspective and, moreover, requires that all of the market microstructure complications inherent in the high–frequency data be properly incorporated into the model, we will make use here of the realized variation measures. Note that the idea to exploit the information contained in the realized variation measures for the estimation of continuous–time stochastic volatility models is not novel to the literature. In fact Bollerslev and Zhou (2002) propose a general method of moment approach, and Barndorff-Nielsen and Shephard (2002a) suggest to use a quasi maximum likelihood. However, both approaches require the derivation of conditional moments of the dynamics of the model–implied realized variation, which is not feasible for all types of continuous–time stochastic volatility models. In contrast, we adopt here the general scientific modeling (GSM) method recently proposed by Gallant and McCulloch (2005), which does not rely on the derivation of such quantities, and allows the assessment of any type of stochastic volatility model (as long as we can simulate from it) within a unified framework. In Chapter 3 we conduct the statistical assessment of univariate continuous–time stochastic volatility models.

Apart from the adequate modeling of the price process of a particular asset, as is the focus of the above mentioned univariate continuous–time stochastic volatility

models, the knowledge of the correlation structure, is also crucial for financial decision-making, such as optimal portfolio choice and asset allocation decisions. In the multivariate context, the model needs not only to capture the individual dynamics, but should also reproduce the comovements and spill-over effects across different assets. As such, modeling becomes even more challenging. Moreover, the multivariate modeling is subject to some technical problems. One is given by the necessity of a positive semidefinite covariance matrix. For stochastic volatility models this implies that the instantaneous covariance matrix should be specified by a positive semidefinite process. Moreover, if the dimension of the return vector increases the number of parameters in the model is inflated. Hence, a parsimonious but at the same time accurate specification is needed. Although the continuous-time specification is very important for the asset pricing perspective, we are aware of only three papers that consider continuous-time multivariate stochastic volatility models, see Hubalek and Nicolato (2005), Lindberg (2005) and Gouriéroux (2006). However, none of these models provide closed-form expressions for the integrated covariance process—the main variable of interest for financial applications.

The fourth chapter of this thesis therefore introduces a new continuous-time multivariate stochastic volatility model that is shown to meet the above mentioned requirements while providing a closed-form and very simple structure for the integrated covariance process. In particular, our model is a multivariate extension of the non-Gaussian Ornstein–Uhlenbeck-type model proposed by Barndorff-Nielsen and Shephard (2001b). As this modeling framework allows us to derive state space representations for the realized covariance matrix and for the squared high-frequency returns, we also assess the adequacy of our multivariate model using high-frequency data. This is in line with the quasi maximum likelihood estimation approach proposed by Barndorff-Nielsen and Shephard (2002a) for the univariate non-Gaussian Ornstein–Uhlenbeck-type stochastic volatility models.

The remainder of this thesis is structured as follows. The next chapter discusses the information contained in high-frequency financial data. In particular, we review the realized variation and covariation measures and illustrate their empirical properties using a univariate and a multivariate dataset, which will be used later on in the empirical assessment of the univariate and multivariate stochastic volatility models, respectively. Chapter 3 presents the statistical assessment of the univariate continuous-time stochastic volatility models. This also involves the derivation of a highly accurate discrete-time model for daily returns and realized variation. The chapter is primarily based on the papers by Bollerslev et al. (2006a) and Bollerslev et al. (2007). Chapter 4 is based on Pigorsch and Stelzer (2007) and introduces the multivariate extension of the non-Gaussian Ornstein–Uhlenbeck-type stochastic volatility model, along with a Monte-Carlo analysis for the assessment of the finite sample properties of the relevant estimation methods. Furthermore, the model is estimated using intraday returns sampled at different frequencies. Chapter 5 concludes.

## 2 High–Frequency Information

With the availability of transaction prices of financial assets, the question arises whether such data provides any new information when compared to the commonly available daily data. As a consequence, a new branch in the financial econometrics literature has developed over the recent years addressing this issue, revealing that the high–frequency data is indeed very informative about the price process not only on an intradaily level—as might be naturally expected—but also on a daily level. In particular, assessing the high–frequency returns directly, i.e. on an intradaily basis, shows mainly that the markets are quite efficient and immediately incorporate news, such as macroeconomic news announcements; that the very highly sampled returns are subject to market microstructure noise induced by the trading mechanism; and that there exists particular intradaily patterns, such as the well-known U–shaped volatility pattern, exhibiting a high volatility at the beginning and at the end of the trading day inducing strong seasonality patterns in the autocorrelation functions of intraday absolute returns, or a high volatility period during lunch (see e.g. Andersen et al., 2003b; Andersen and Bollerslev, 1997; Bollerslev et al., 2000; Engle and Russell, 2007). However, all of these effects complicate the direct modeling of high–frequency returns and in contrast to the daily level, for which GARCH–type models are widely accepted as a quite accurate description of the daily returns, a similar unifying and adequate approach accounting for all of these intraday specific effects has not been established yet. Instead, rather than modeling the raw high–frequency returns directly, the returns are usually adjusted for some or all of these effects.

Alternatively, the information contained in the high–frequency data can be summarized on a lower frequency, usually the daily level, resulting in the so–called realized variation measures. In particular, based on the theory of quadratic variation the sum over the outer product of high–frequency returns provides an ex–post measure of the daily quadratic (co)variation—a key variable in many financial applications. Moreover, these measures provide new information on the daily volatility dynamics and the distribution of the volatility and standardized returns. They also allow to empirically distinguish between the price variation coming from the continuous–time evolution of the price process and the variation coming from jumps, and as such are informative on the contribution of jumps to total price variation. Furthermore, the relationships between the two volatility components and returns can be assessed.

As a consequence, the nonparametric volatility measures have lead to the development of a series of new and simple–to–implement reduced form volatility forecasting models in which the realized volatilities are modeled by standard discrete–time

time series procedures, examples of which include Andersen et al. (2003a, 2007), Corsi (2004), Corsi et al. (2007), Deo et al. (2006), Koopman et al. (2005) and Martens et al. (2004), among others. Noteworthy, by effectively incorporating the high-frequency data into the volatility measurements, these simple discrete-time models generally out-perform existing more complicated parametric volatility models based on the corresponding return observations only, such as the GARCH-type models, being indicative of the higher information content of these measures when compared to the daily returns.

Given the usefulness of the realized variation measures for modeling volatility, we first review the theory of quadratic variation in the multivariate setup, and then provide the definitions of realized variation, Bipower variation and a jump measure for the special case of a univariate price process, see section 2.1.1. Thereafter, we discuss the realized covariation, see section 2.1.2. We also establish some important notation. In the sequel, we provide a discussion of the univariate dataset used for the statistical assessment of the univariate continuous-time stochastic volatility models (Section 2.2.1), as well as of the multivariate dataset used in the empirical application of our newly developed multivariate extension of the OU-type stochastic volatility model (Section 2.2.2). Within each of these sections we also exemplify the empirical properties of the resulting series.

## 2.1 Definition of Realized Variation and Covariation Measures

Let  $Y_t$  denote the logarithmic price process of  $d$  different financial assets, and assume that it belongs to the following class of stochastic volatility semimartingales:

**Definition 2.1.1.** *A stochastic volatility semimartingale denoted by  $\mathcal{SVSM}$  is a vector semimartingale  $Y = \alpha + m$  satisfying the following conditions:*

- (i) *that  $\alpha \in \mathcal{FV}_{loc}^c$ , i.e. the drift has locally finite and continuous sample paths, and  $\alpha(0) = 0$ .*
- (ii) *that  $m$ , the multivariate stochastic volatility process, is a local martingale.*

Note, that our definition is quite similar to that of Barndorff-Nielsen and Shephard (2004a), however, we allow the local martingale component of the logarithmic prices in (ii) to have jumps. Moreover, even if the volatility process exhibits jumps,  $Y$  is still a stochastic volatility semimartingale. Given  $Y \in \mathcal{SVSM}$ , the *quadratic variation* or *covariation* process, generally defined as (see e.g. Jacod and Shiryaev,

2003)<sup>1</sup>

$$[Y]_t := \text{plim}_{M \rightarrow \infty} \sum_{j=1}^M (Y_{t_j} - Y_{t_{j-1}})(Y_{t_j} - Y_{t_{j-1}})^T, \quad (2.1)$$

for any sequence of partitions  $t_0 = 0 \leq t_1 \leq \dots \leq t_M = t$  with  $\sup_j t_j - t_{j-1} \rightarrow 0$  for  $M \rightarrow \infty$ , is given for  $Y$  as

$$[Y]_t = [Y^c]_t + [Y^d]_t \quad (2.2)$$

$$= [Y^c]_t + \sum_{0 \leq s \leq t} (\Delta Y_s)(\Delta Y_s)^T, \quad (2.3)$$

where  $[Y^c]$  and  $[Y^d]$  are the continuous and the discontinuous local martingale (or quadratic variation) components, respectively, whereby  $\Delta Y_t = Y_t - Y_{-t}$  denote the jumps occurring at time  $t$ . Furthermore, since (i) holds,

$$[Y]_t = [m]_t = [m^c]_t + \sum_{0 \leq s \leq t} (\Delta m_s)(\Delta m_s)^T, \quad (2.4)$$

with  $\Delta m_t := m_t - m_{-t}$ , see e.g. Barndorff-Nielsen and Shephard (2004b).

For financial applications, the knowledge of the total price variation process, and of its two components is essential and thus deriving consistent measures of these quantities is important. The theory of quadratic variation, more specifically equation (2.1), suggests that summing over the outer product of ideally infinitesimally sampled return vectors can provide an ex-post empirical measure of the quadratic variation at time  $t$  if the sum is computed over the time interval  $[0, t]$ . In a similar manner, the quadratic variation can be measured over any other time interval as long as the summation interval is adjusted correspondingly. With the availability of high-frequency data such an approach has become feasible. In particular, as already shown by Merton (1980) and extended by Andersen and Bollerslev (1998), Andersen et al. (2001b), Barndorff-Nielsen and Shephard (2001b), and by Comte and Renault (1998), the quadratic variation can indeed be consistently estimated by the sum of squared returns computed over very small time intervals. These results hold even if the exact form of the drift and volatility processes are unknown (see Barndorff-Nielsen and Shephard, 2002a). Although these authors are the first to establish the formal relationship of the notion of realized variation to the theory of quadratic variation within the context of finance and time-varying volatility modeling, the idea of measuring the ex-post variation of asset prices by summing over more frequently sampled squared returns dates back at least to Merton (1980), and was also applied by French et al. (1987), Hsieh (1991) and Poterba and Summers (1986), and more recently by Taylor and Xu (1997), inter alia. Moreover, based on the theoretical results derived in Barndorff-Nielsen and Shephard (2004b) it is also possible to construct measures of the two volatility components.

---

<sup>1</sup>In the following  $^T$  denotes the transposed vector or matrix. As is common practice, all vectors are column vectors.

## 2 High-Frequency Information

Before discussing these realized quadratic (co)variation measures, we first introduce some notation. Generally, we are interested in the discretely observed logarithmic price increments of  $Y_t$  over unit time intervals of length  $\Delta$ . We therefore denote the return over the time interval  $[(n-1)\Delta, n\Delta]$  with  $n \in \mathbb{N}$  by  $\mathbf{Y}_n$ , i.e.

$$\mathbf{Y}_n := Y_{n\Delta} - Y_{(n-1)\Delta}. \quad (2.5)$$

As is commonly done, we thereby focus on the *daily returns*, i.e. in (2.5)  $\Delta$  represents one day.

In addition, for each  $n$  (e.g. for each day) we observe the logarithmic price increments over subintervals of  $\Delta$ . In particular, the *high-frequency returns* (or intradaily returns) are denoted by

$$\mathbf{Y}_{j,n} := Y_{((n-1)+\frac{j}{M})\Delta} - Y_{((n-1)+\frac{j-1}{M})\Delta} \quad \text{with } j = 1, \dots, M, \quad (2.6)$$

whereby  $M$  refers to the sampling frequency. Commonly, the length of the subintervals  $\Delta/M$  is e.g. 5, 15, or 30 minutes.

In the next section we discuss the notion of realized variation, Bipower variation and the logarithmic jump measure within the context of a univariate price process. Thereafter, the realized covariation measure is reviewed.

### 2.1.1 Realized Variation

Assume that for the univariate price process  $y_t \in \mathcal{SVSM}$ , and that the stochastic volatility process is given by the following Brownian semimartingale plus jumps

$$m_t = \int_0^t \sigma(s) dw(s) + \sum_{j=1}^{N(t)} \kappa(s_j), \quad (2.7)$$

where  $\sigma(t) > 0 \quad \forall t$  denotes the càdlàg instantaneous stochastic volatility,  $w(t)$  is a standard Brownian motion, and the  $N(t)$  process counts the (for all  $t$  assumed finite) number of jumps occurring with possibly time-varying intensity  $\lambda(t)$  and jump size  $\kappa(s_j)$ . Note that most of the commonly used continuous-time stochastic volatility models are subsumed in this class, i.e. the logarithmic and affine jump-diffusion models or the non-Gaussian Ornstein-Uhlenbeck type processes proposed by Barndorff-Nielsen and Shephard (2001b) and their extensions. The theory of quadratic variation then allows to derive nonparametric realized variation measures that can be used to decompose the total price variation into the variation coming from the continuous sample path evolution and the variation coming from the jumps. In particular, for (2.7) the quadratic variation process is given by

$$[y]_t = [y^c]_t + [y^d]_t = \int_0^t \sigma^2(s) ds + \sum_{j=1}^{N(t)} \kappa^2(s_j), \quad (2.8)$$

that is, the quadratic variation is the *integrated variance*, i.e. the continuous local martingale or quadratic variation component, plus the sum of the *squared jumps*. Obviously, in the familiar pure diffusion case where the  $N(t)$  counting process is identically equal to zero, i.e.  $[y]_t = [y^c]_t$ , the second term disappears and the quadratic variation is simply equal to the integrated variance.

Then, by the theory of quadratic variation of semimartingales the *realized variance* over day  $n$  defined by

$$[y]_n^{(M)} := \sum_{j=1}^M \mathbf{Y}_{j,n}^2, \quad (2.9)$$

whereby the superscript  $M$  indicates the dependence of this quantity on the sampling frequency, converges uniformly in probability to the (daily) quadratic variation process as the sampling frequency of returns approaches infinity, i.e., for  $M \rightarrow \infty$

$$[y]_n^{(M)} \xrightarrow{P} \int_{(n-1)\Delta}^{n\Delta} \sigma^2(s) ds + \sum_{j=N(n-1)+1}^{N(n)} \kappa^2(s_j). \quad (2.10)$$

In other words, the realized variance provides a consistent measure of the true total price variation, including the discontinuous jump part.

In order to distinguish the continuous variation from the jump component, Barndorff-Nielsen and Shephard (2004b) first proposed the so-called *Bipower variation* measure, defined by

$$[y]_n^{1,1(M)} := \frac{\pi}{2} \sum_{j=2}^M |\mathbf{Y}_{j,n}| |\mathbf{Y}_{j,n}|, \quad (2.11)$$

whereby we have basically adopted their notation.<sup>2</sup> Importantly, for increasingly finely sampled returns the Bipower variation measure becomes robust to jumps and consistently (for increasing values of  $M$ ) estimates the integrated variance

$$[y]_n^{1,1(M)} \xrightarrow{P} \int_{(n-1)\Delta}^{n\Delta} \sigma^2(s) ds. \quad (2.12)$$

Consequently, the difference between the realized variance and the Bipower variation provides a consistent nonparametric estimator of the pure jump contribution to total price variation, and can be used for testing for the presence of jumps as advocated empirically by Andersen et al. (2007) using the theoretical results derived in Barndorff-Nielsen and Shephard (2006a), who consider a ratio jump statistics. Alternatively, we consider the logarithmic relative jump measure defined by

$$\{y\}_n^{d(M)} = \log[y]_n^{(M)} - \log[y]_n^{1,1(M)}, \quad (2.13)$$

---

<sup>2</sup>In particular, the superscript 1,1 refers for the powers of the current and lagged absolute intraday returns.



for which the corresponding test statistic might possess better finite sample performance given the results in Barndorff-Nielsen and Shephard (2005) showing that the variance of the logarithmic realized variance estimator is smaller than for the non-transformed realized variance.

### 2.1.2 Realized Covariation

Our multivariate stochastic volatility model that is introduced in Section 4 is an extension of the univariate non-Gaussian Ornstein–Uhlenbeck–type models proposed by Barndorff-Nielsen and Shephard (2001b, 2002a). Since these models are inter alia characterized by a pure diffusion price process, we constrain ourselves on the discussion of the realized covariation for the class of continuous stochastic volatility semimartingales, i.e.  $Y_t \in \mathcal{SVSM}^c \in \mathcal{SVSM}$ . In particular, we assume that

$$m_t = m_t^c = \int_0^t \Sigma^{1/2}(s) dW(s), \quad (2.14)$$

whereby  $\Sigma(s)$  is the instantaneous covariance process with values in the positive semidefinite matrices and càdlàg elements, and  $W(t)$  denotes a  $d$ -dimensional standard Brownian motion. In this case, the quadratic covariation is given by

$$[Y_t] = \Sigma_t^+ := \int_0^t \Sigma_t dt, \quad (2.15)$$

i.e. the integrated covariance matrix, which is of main interest for financial applications. Oftentimes, we are interested in the daily integrated covariance, which can be measured ex-post by the so-called *realized covariation matrix* using high-frequency returns

$$[Y]_n^{(M)} := \sum_{j=1}^M \mathbf{Y}_{j,n} \mathbf{Y}_{j,n}^T, \quad (2.16)$$

whereby the superscript  $(M)$  reflects the dependence of this measure upon the particular sampling frequency  $M$ . Generally, by the theory of quadratic variation it follows, that for  $M \rightarrow \infty$

$$[Y]_n^{(M)} \xrightarrow{p} [Y]_n - [Y]_{n-1}, \quad (2.17)$$

i.e. realized covariation is a consistent estimator of the daily increment of quadratic variation. Moreover, if the log-price process  $Y_t \in \mathcal{SVSM}^c$ , realized covariation consistently estimates the integrated covariance over day  $n$  for  $M \rightarrow \infty$ ,

$$[Y]_n^{(M)} \xrightarrow{p} \Sigma_n, \quad (2.18)$$

with

$$\Sigma_n := \Sigma_{n\Delta}^+ - \Sigma_{(n-1)\Delta}^+ = \int_{(n-1)\Delta}^{n\Delta} \Sigma_t dt, \quad (2.19)$$

as shown in Barndorff-Nielsen and Shephard (2004a), who further derive the asymptotic theory of realized covariance as an estimator of the increments of quadratic variation or integrated variance.



## 2.2 Stylized Facts of Returns and Realized Variation Measures

This section describes the datasets we use in our univariate and multivariate model assessment, along with the empirical properties of the different realized variation measures. A discussion of the issues arising in the practical computation of these measures is also provided.

### 2.2.1 Univariate Dataset

The analysis of the univariate stochastic volatility models is based on high-frequency S&P500 index futures data. In particular, our dataset consists of tick-by-tick transaction prices of S&P500 index futures recorded at the Chicago Mercantile Exchange (CME). The sample ranges from January 1, 1985 to December 31, 2004, a period of 5,040 trading days with 13,241,032 tick-by-tick observations. In the following we discuss the construction of the daily realized variation, Bipower variation and the logarithmic relative jumps measures.

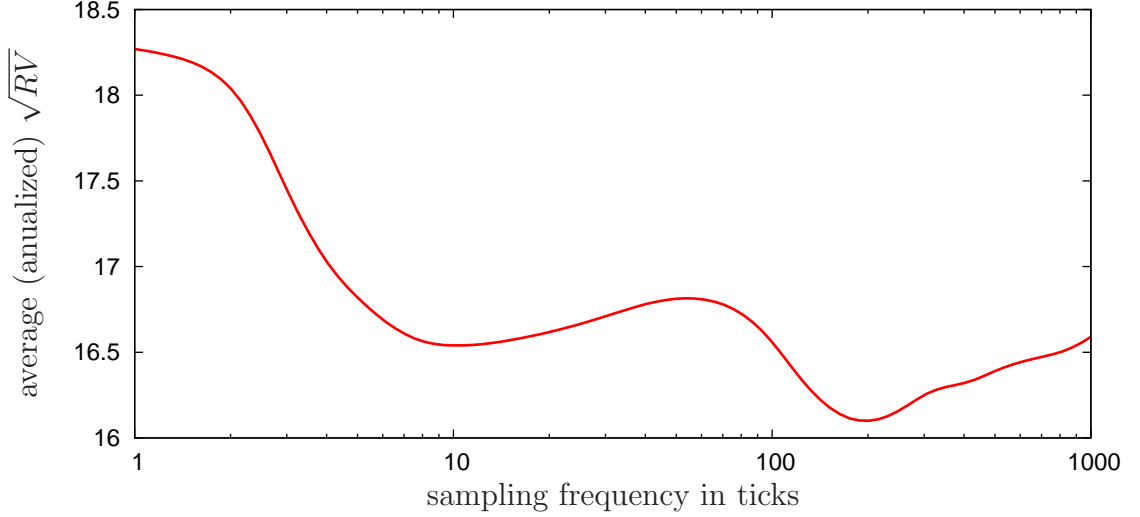
### Construction of Realized Variation Measures

It follows from the theoretical considerations discussed above that the consistency of the realized variation, Bipower variation and the logarithmic relative jumps (as well as for the realized covariation measure) hinges on the notion of increasingly finer sampled high-frequency returns. In practice, however, the sampling frequency is invariably limited by the actual quotation, or transaction frequency. Moreover, very high-frequency returns, e.g. computed over 1 minute or even shorter time intervals, are contaminated by transaction costs, bid-and-ask-bounce effects etc., leading to biases in the variance measures and rendering the basic assumption of a semimartingale price process to be invalid at the very high-frequency level. Consequently, the existence of such market microstructure noise induces a bias-variance trade-off when constructing the respective measures (see also Bandi and Russell, 2005a; Zhang et al., 2005). In particular, in order to achieve consistency, a high sampling frequency is required leading to the accumulation of market microstructure noise with the bias tending to become more severe as the sampling frequency increases. On the other hand, using lower frequencies will result in less precise estimates. In response to this, a number of authors, including Andersen et al. (2001a,b, 2007), have advocated the use of coarser sampling frequencies, such as 5 to 30 minutes as a simple way to alleviate the contaminating effects, while maintaining most of the relevant information in the high-frequency data. Alternatively, different procedures have been proposed in the literature that make use of the very high-frequency returns, e.g. computed even on a tick-by-tick basis. Since the market microstructure noise induces autocorrelation in the intraday returns which in turn leads to the bias problem, these approaches adopt techniques that are usually

applied in the estimation of the variance of a stationary time series in the presence of autocorrelation (see also Hansen and Lunde, 2006). For example, for the estimation of realized variance, pre-whitening techniques, such as the moving-average filter (see e.g. Andersen et al., 2001a; Hansen et al., 2007) or the autoregressive filter of Bollen and Inder (2002), nonparametric techniques, such as kernel-based estimators (see Hansen and Lunde, 2006; Zhou, 1996) or estimators based on subsampling (e.g. Zhang et al., 2005; Zhou, 1996) have been proposed. The recent paper by Barndorff-Nielsen et al. (2006a) and Barndorff-Nielsen et al. (2006b) provides a unified theoretical framework for analyzing most of these estimators within a kernel-based representation along with a discussion of optimal kernel and bandwidth choices. In particular, they derive conditions under which these estimators are consistent and very close to efficient. Moreover, their results are robust to rather broad assumptions about the market microstructure noise dynamics. However, the asymptotic theory is derived under the assumption of a pure Brownian semimartingale, i.e. there are no jumps in the price process. Other approaches build on the notion of an optimal sampling frequency,  $M$ , in the sense of minimizing the MSE of the resulting realized volatility measure as suggested by Bandi and Russell (2005a) and Aït-Sahalia et al. (2005), or of business type sampling schemes dictated by the activity of the market, as in, e.g., Oomen (2005).

So far, the literature has focused only on bias-correcting the realized variance measure, and to the best of our knowledge none of these ideas have yet been formally extended to allow for similar measurements of the integrated variance in form of robust to market microstructure noise modified realized Bipower variation. Consequently, similar work for the construction of market microstructure robust jump measures and jump detection tests is still pending. Being interested in the decomposition of total price variation into its two components, we therefore address the bias-variance trade-off by sampling at lower frequencies. In particular, given the high liquidity of our S&P500 index futures data, we follow Andersen and Bollerslev (1998), Andersen et al. (2001b), Maheu and McCurdy (2002) and Martens et al. (2004), among others, and use five-minute returns to construct our realized-variance, Bipower-variation and jump measures.

The computation of the realized-variation measures is based on the most liquid contracts. In particular, we consider the transaction prices of the most liquid contract at the beginning of our sample period and switch to another contract if this is traded more frequently. The corresponding intraday returns are then constructed from the transaction prices of each of these contracts, i.e. we avoid to compute returns over the roll-over period. Moreover, we use the nearest neighbor to the five-minute mark and exclude overnight returns, since the overnight trading of these contracts at GLOBEX—the CME overnight trading platform—just started in 1994. Using the same dataset and construction methods, Corsi et al. (2007) have shown that the impact of market microstructure noise on these realized-variation measures is negligible. In particular, they make use of the volatility-signature plot which is a useful tool for assessing the bias induced by the microstructure noise by depicting the full sample averages of realized volatility computed for different



**Figure 2.1:** Volatility–signature plot of the S&P500 index futures constructed over the full sample period. The graph shows average annualized realized volatility constructed for different frequencies measured in number of ticks. Note that there are about 7 seconds on average between trades, such that the average annualized five–minute based realized volatility corresponds to around the 43th tick.

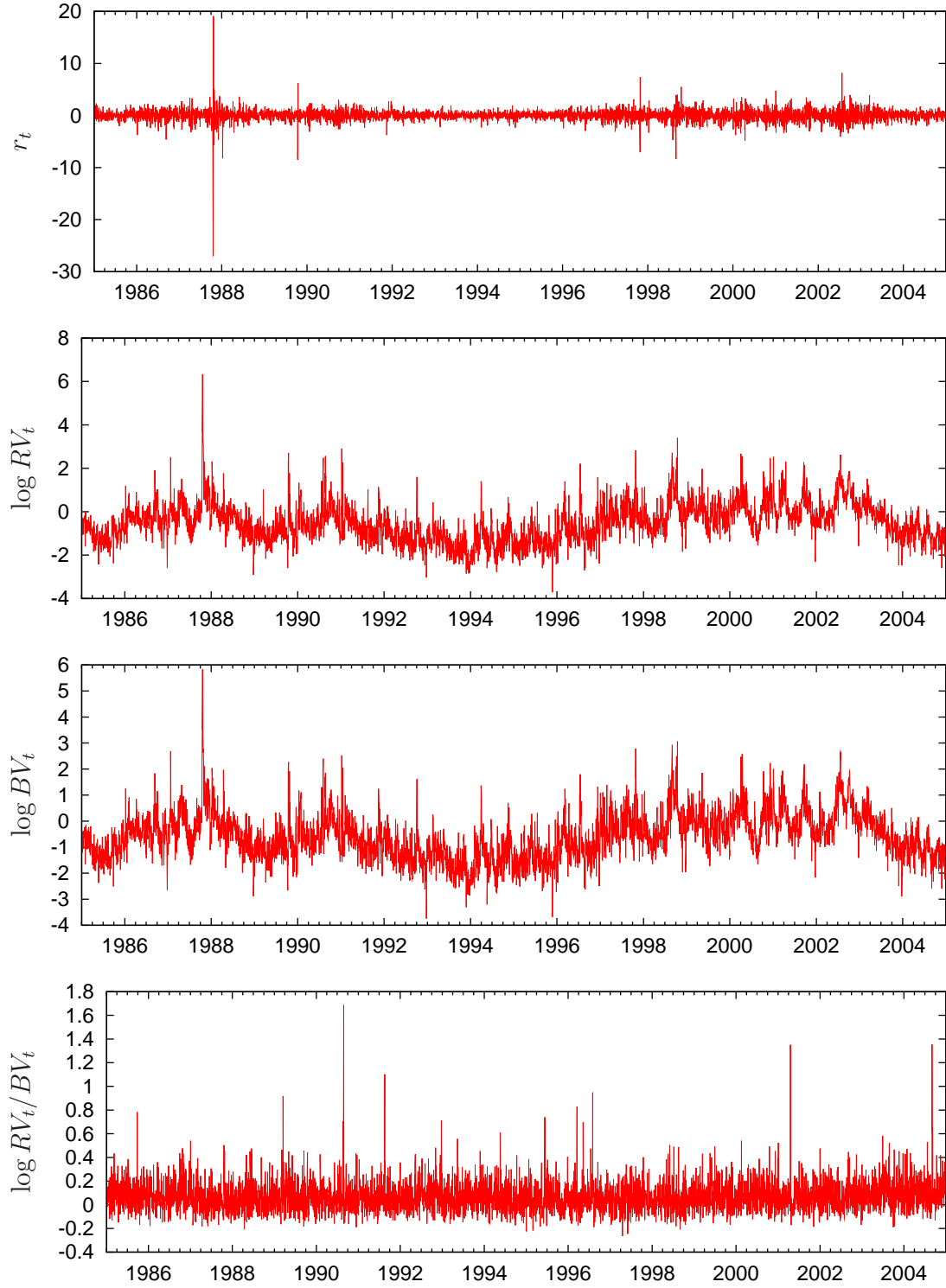
frequencies. In Figure 2.1 we reproduce the volatility–signature plot of Corsi et al. (2007) for the S&P500 index futures. Note that a transaction takes place on average about every seven seconds, such that the average annualized realized volatility based on the five–minute intervals corresponds to around the 43th tick presented in the Figure. Obviously, the bias dies out very quickly.

### Empirical Properties

In the following we discuss and illustrate the empirical properties of the resulting daily returns, realized variation, Bipower variation and jump measures. Note that we basically reproduce here the descriptive data analysis of Bollerslev et al. (2007). Moreover, for the ease of exposition we denote from now on the realized variance by  $RV_t := [y]_n^{(M)}$ , the Bipower variation by  $BV_t := [y]_n^{1,1(M)}$  and the logarithmic relative jump measure by  $J_t := \{y\}_n^{d(M)}$ .<sup>3</sup>

The daily return, logarithmic realized–variance, logarithmic Bipower–variation

<sup>3</sup>The somehow less intuitive notation in Section 2.1.1 was chosen to provide a unified framework for discussing the univariate as well as the multivariate measures. Moreover, the notation was partly adopted from Barndorff-Nielsen and Shephard (2004a) and turns out to be useful in the derivation of our multivariate stochastic volatility model. In contrast, the univariate model assessment relies on the existence of a discrete–time model for these measures, and such notation would complicate the intuitive representation of this auxiliary model.



**Figure 2.2:** Time Series of returns, logarithmic realized variance, logarithmic Bipower variation and jumps.

**Table 2.1:** Descriptive Statistics of the Univariate Dataset

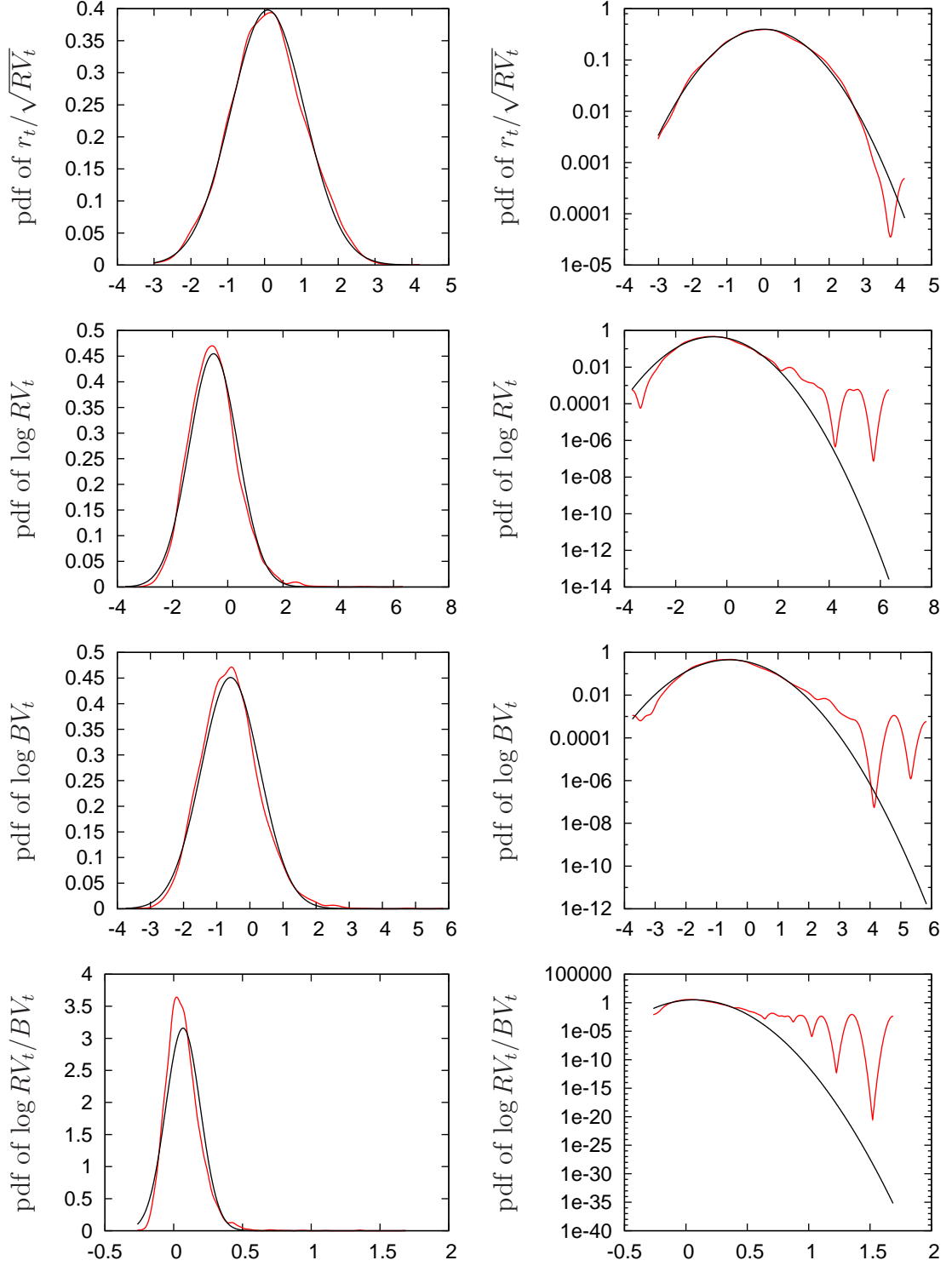
Series	Mean	Std.Dev.	Median	Skewness	Exc.Kurt.	Ljung-Box(10)
$\sqrt{RV_t}$	0.8627	0.5935	0.7586	15.3509	496.7651	10155.72
$\log RV_t$	-0.5139	0.8775	-0.5527	0.5950	1.7981	22023.20
$\sqrt{BV_t}$	0.8340	0.5359	0.7348	11.1561	288.4633	12223.28
$\log BV_t$	-0.5817	0.8845	-0.6163	0.5418	1.4807	21715.55
$\log \left( \frac{RV_t}{BV_t} \right)$	0.0678	0.1263	0.0538	1.7766	12.2675	51.44
$r_t$	0.0254	1.0946	0.0511	-2.1655	96.2483	117.29
$r_t/\sqrt{RV_t}$	0.0866	1.0027	0.0739	0.0503	-0.1497	14.86

and jump series are displayed in Figure 2.2. The widely-documented volatility-clustering effect becomes obvious in all three series. Moreover, the realized-variance is more volatile than the Bipower-variation series, which might be due to the jump series exhibiting many, mostly positive, small values. Some of these observations, and of the small negative values, may be attributed to measurement or discretization errors induced by sampling at a lower frequency in order to eliminate the market microstructure bias. But there are also larger values, which, in contrast, can be associated with genuine large-sized jumps on those particular days. Although we do not test for the presence of jumps here, these observations are indicative of a relevant contribution of jumps to total price variation. In fact, using similar data, Huang and Tauchen (2005) find that jumps contribute to total price variation by about 7%.

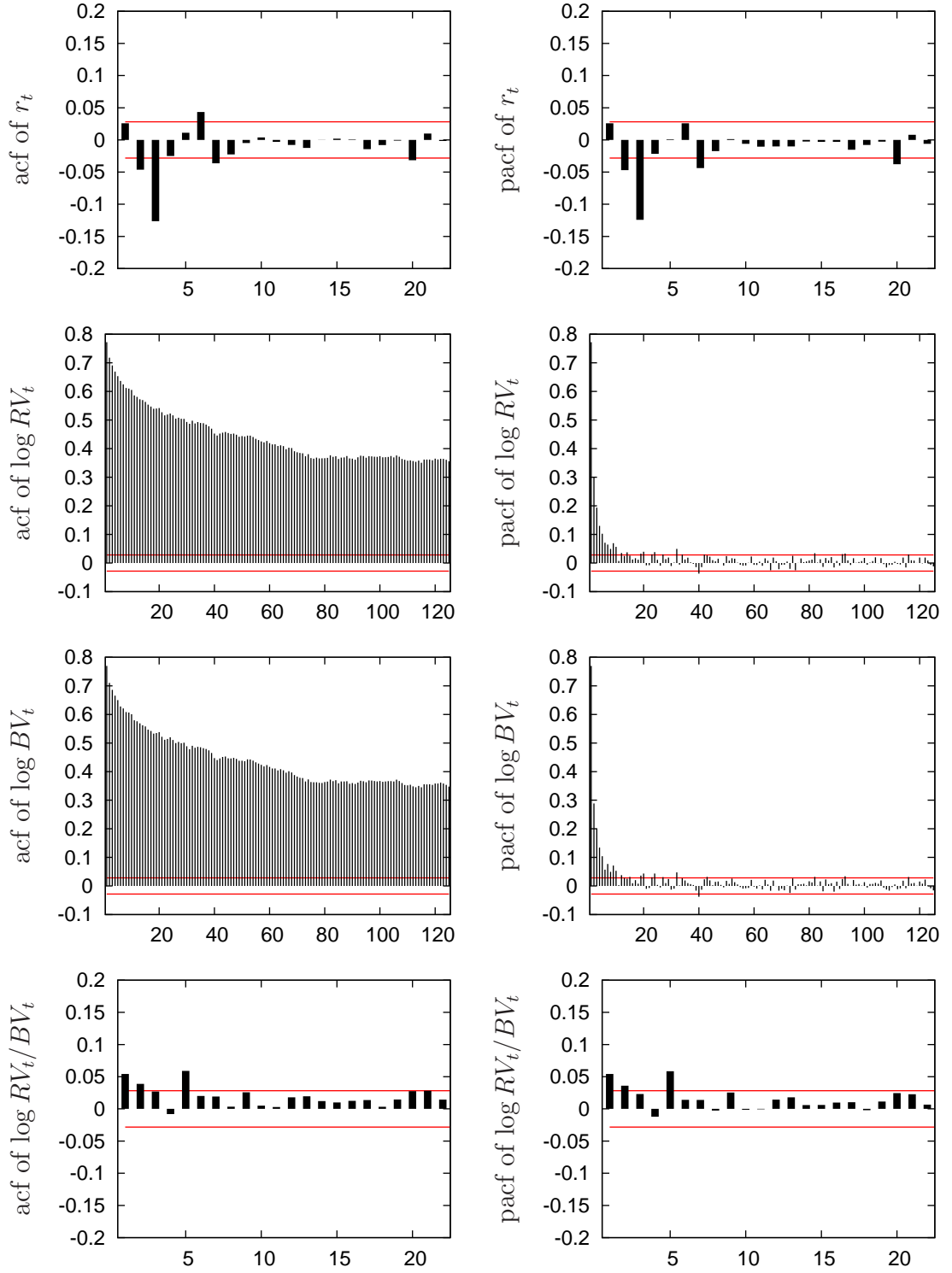
The visual impressions are confirmed by the summary statistics reported in Table 2.1. In particular, the mean and variance of the realized volatility exceed the corresponding statistics for the square-root Bipower variation. Also, the unconditional distribution of both volatility measures is highly skewed and leptokurtic, but can be made close to Gaussianity by the logarithmic transform, which is further supported by the kernel density plots presented in Figure 2.3, and is in line with the empirical findings in Andersen et al. (2001a,b), among others. The descriptive statistics and the corresponding kernel density plots for the relative jump measure,  $J_t$ , clearly indicate a positively skewed and leptokurtic distribution.<sup>4</sup> The unconditional distribution of the daily returns also shows the well-known excess kurtosis and negative skewness, and is surprisingly close to Gaussianity if the distribution is scaled by the realized volatility, as previously documented by Andersen et al. (2001a).

According to the Ljung-Box test statistics for up to tenth order autocorrelation,

<sup>4</sup>Note that the sign of the skewness is determined by the specific definition of our jump measure as the ratio of  $RV_t$  divided by  $BV_t$ . Barndorff-Nielsen and Shephard (2004b) in contrast consider the inverse ratio resulting in a negatively skewed distribution.



**Figure 2.3:** Unconditional distributions of standardized returns, logarithmic realized variance, logarithmic Bipower variation and jumps. The left panel of the figure shows the kernel density estimates of the series (red line) and the normal density (black line) for reference purposes. The right panel shows the same in log scale.



**Figure 2.4:** Sample autocorrelations and partial autocorrelations of returns, logarithmic realized variance, logarithmic Bipower variation and jumps. The red lines give the upper and lower ranges of the conventional Bartlett 95% confidence band.

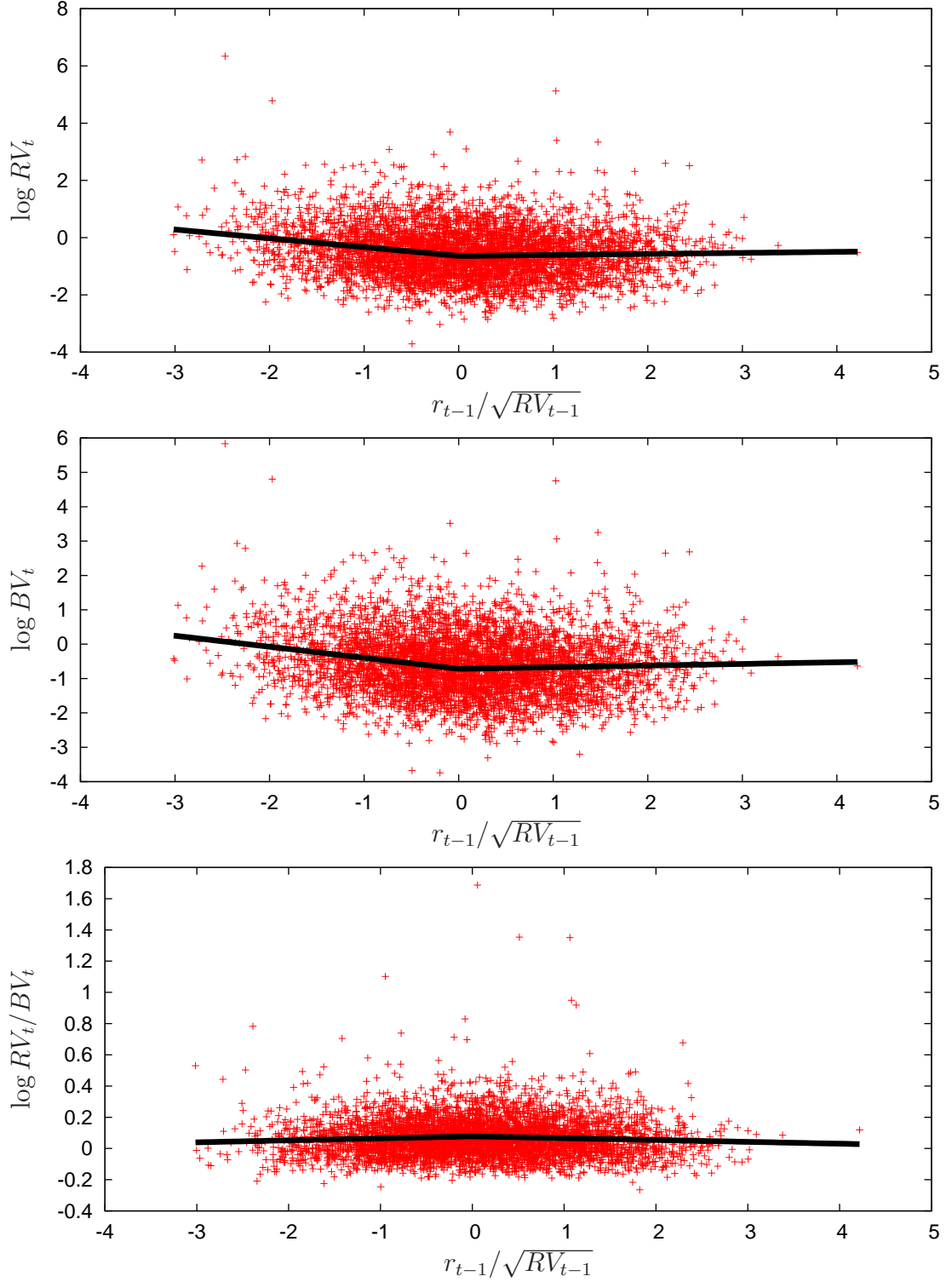
all of the volatility measures exhibit highly significant own serial dependencies. This is also supported by the sample autocorrelation functions presented in Figure 2.4 showing significant autocorrelation coefficients (compared to the conservative Bartlett 95% confidence bands) up to the 125th order corresponding to roughly half-a-year. The Figure also visualizes the nearly hyperbolic decay of the autocorrelation functions for the two logarithmic volatility measures—a characteristics that has also been reported in the GARCH and stochastic volatility literature. In contrast, the relative jump measure exhibits much less autocorrelation, with most of the dependency being attributable to the first and the fifth lag, corresponding to jumps that are one day and one week apart, respectively. Such weak autocorrelation in the jump series has also been found in Andersen et al. (2007).

Another stylized fact, that is also well-known from the GARCH- and stochastic-volatility literature, is the negative correlation between past return shocks and current volatility, in particular negative return shocks tend to be associated with a larger increase in volatility than a positive return shock of the same magnitude. This phenomenon is very often referred to as the leverage effect, although for equity indices the observed effect is too large to be caused by financial leverage, and such an explanation would be more adequate for single equities. Alternatively, the effect is sometimes also explained by the existence of a time-varying risk-premium implying that expected returns depend positively on the conditional volatility, see for example Bekaert and Wu (2000), Campbell and Hentschel (1992) as well as French et al. (1987). Obviously, the two explanations assume a converse causality and the empirical evidence on both effects is controversial. However, the recent high-frequency data analysis in Bollerslev et al. (2006b) points towards a "leverage"-type causality. The news-impact curve proposed by Engle and Ng (1993) is a common approach to empirically visualize this asymmetric relationship. The corresponding plots for the logarithmic realized variance and Bipower variation are given in Figure 2.5. Both exhibit the expected slight asymmetric response to past standardized returns. Interestingly, however, such relationship is not found for the jumps which seem to be almost unaffected by the past return shocks, and, if anything, respond negatively to the standardized returns. This also explains, why the asymmetric effect is more pronounced for the pure continuous volatility  $BV_t$  component in the second panel, as compared to the total realized variation  $RV_t$  depicted in the first panel.

### 2.2.2 Multivariate Dataset

Our multivariate analysis is based on tick-by-tick transaction prices of various US stocks. Table 2.2 provides some information on the companies included in our dataset. As can be seen, our study includes companies of very different size and from different sectors. The data is taken from the Trade and Quote (TAQ) Database and covers the period from January 1, 2001 to December 31, 2005, a period of 1,256 trading days. The descriptive statistics are presented in Table 2.3, whereby we have made the following data adjustments. We only use transactions





**Figure 2.5:** News-impact curves for logarithmic realized variance, logarithmic Bipower variation and jumps. The figure shows the scatter points between the respective variable and lagged standardized returns. The black lines are the news-impact curves, i.e. the linear regression lines for negative and positive values of standardized returns.

**Table 2.2:** Company Descriptions of the Multivariate Dataset

Symbol	Name	Sector	Employees
AA	Alcoa Inc.	Basic Materials	129,000
C	Citigroup Inc.	Financial	299,000
HAS	Hasbro Inc.	Consumer Goods	5,900
HDI	Harley-Davidson Inc.	Consumer Goods	9,700
INTC	Intel Corp.	Technology	99,900
MSFT	Microsoft Corp.	Technology	61,000
NKE	Nike Inc.	Consumer Goods	28,000
PFE	Pfizer Inc.	Healthcare	106,000
TEK	Tektronix Inc.	Technology	4,359
XOM	Exxon Mobil Corp.	Basic Materials	106,000

taking place during the official trading time, i.e. from 9.30 a.m. to 4 p.m. and exclude overnight returns. Moreover, we only consider valid trades, i.e. we remove trades indicated by “exclude” and “error” flag provided by the TAQ database. The resulting number of total as well as effectively used trades are reported in the second panel of Table 2.3. The lower panel informs about the distribution of the used trades across the different exchanges. The upper panel presents the mean and standard deviation of the daily returns, as well as the duration between trades, i.e. the mean time between used trades measured in seconds. Obviously, our sample consists of quite actively traded stocks with the largest duration being somewhat lower than half a minute. The intraday returns are constructed using the nearest neighbor prior to the corresponding time mark, e.g. prior to the 15 minute tag.

On an intradaily basis, we also find the well-known U-shaped volatility pattern, which is illustrated in Figure 2.6 showing the average number of trades taking place within each minute of a trading day for the four most frequently traded assets of our sample, i.e. for Intel Corp. (INTC), Citigroup Inc. (C), Microsoft Corp. (MSFT) and Pfizer Inc. (PFE). The effect also induces a sinusoidal behavior in the autocovariance of the squared returns as is illustrated by the red line in the upper and lower panels of Figure 2.7, depicting the autocovariance functions for Microsoft Corp. and Intel Corp. To account for this intraday pattern we adjust our dataset by computing the intraday returns of each single stock as

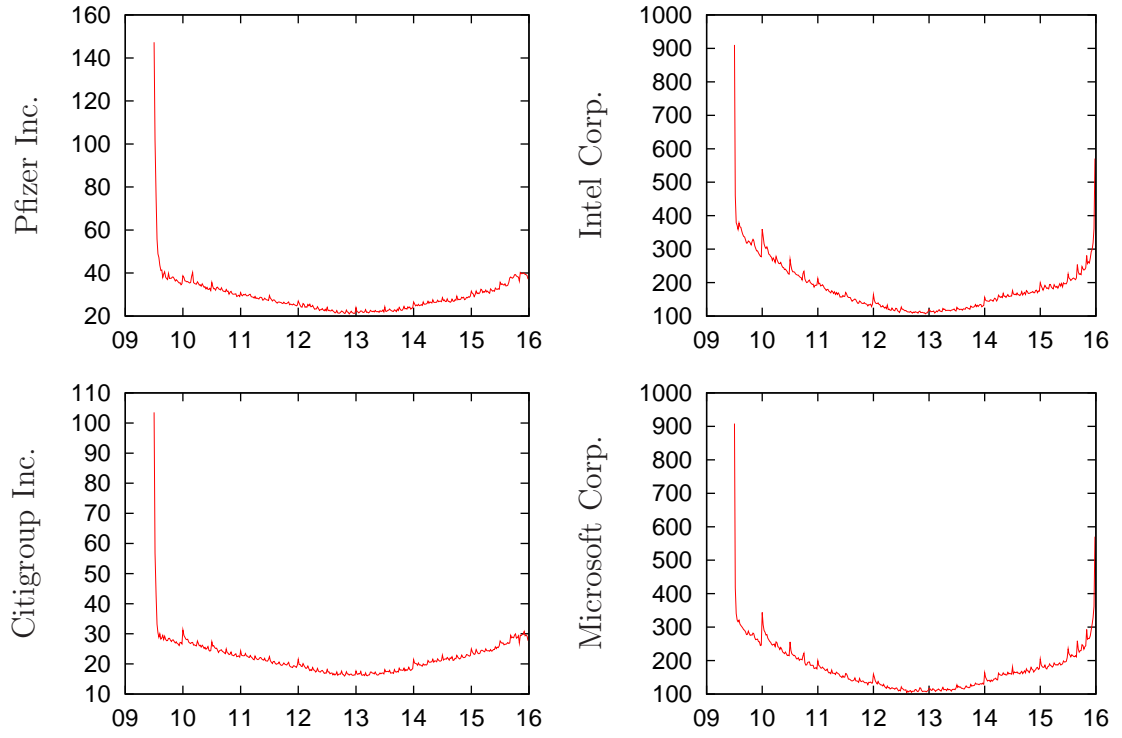
$$y_{j,n}^{(a)} = \frac{\bar{\sigma}}{\sigma_j} y_{j,n} \quad \text{for } j = 1, \dots, M \quad (2.20)$$

whereby  $\bar{\sigma}$  denotes the variance of the intraday returns over the whole sample period and  $\sigma_j$  is the variance over the time interval of length  $\Delta/M$ , i.e. 15 minutes. The black lines in Figure 2.7 show that our data adjustment procedure indeed removes the sinusoidal behavior in the autocovariance functions of the squared intraday returns of MSFT and INTC as well as of their crossproduct. In our empirical application, however, we consider both, the intradaily unadjusted as well as the

**Table 2.3:** Description of the Multivariate Dataset

		AA	C	HAS	HDI	INTC	MSFT	NKE	PFE	TEK	XOM
stats	mean	-0.12	-0.04	0.05	-0.03	-0.04	0.01	0.05	-0.10	-0.07	-0.01
	std	1.83	1.62	1.75	1.82	2.39	1.73	1.56	1.52	2.17	1.27
	dur	6.24	2.67	21.89	8.60	0.32	0.33	10.56	2.02	26.14	2.75
trades	all	4720081	11078678	1344353	3421089	92662950	88698413	2773959	14644488	1121288	10720372
	del	55579 [1.17%]	165524 [1.49%]	13962 [1.03%]	35079 [1.02%]	1965582 [2.12%]	1690277 [1.90%]	24495 [0.88%]	252158 [1.72%]	11285 [1.00%]	139159 [1.29%]
	used	4664502	10913154	1330391	3386010	90697368	87008136	2749464	14392330	1110003	10581213
exchange	A	0 [0.00%]	0 [0.00%]	0 [0.00%]	0 [0.00%]	4377 [0.00%]	24256 [0.02%]	0 [0.00%]	0 [0.00%]	0 [0.00%]	0 [0.00%]
	B	213832 [4.58%]	1004835 [9.20%]	6284 [0.47%]	93700 [2.76%]	2045577 [2.25%]	1756565 [2.01%]	14416 [0.52%]	1685616 [11.71%]	1499 [0.13%]	634833 [5.99%]
	C	46613 [0.99%]	379611 [3.47%]	809 [0.06%]	11665 [0.34%]	22241752 [24.52%]	20222366 [23.24%]	16327 [0.59%]	536066 [3.72%]	151 [0.01%]	289643 [2.73%]
	D	0 [0.00%]	0 [0.00%]	0 [0.00%]	0 [0.00%]	2292952 [2.52%]	2122142 [2.43%]	0 [0.00%]	0 [0.00%]	0 [0.00%]	0 [0.00%]
	M	137632 [2.95%]	386381 [3.54%]	21763 [1.63%]	154569 [4.56%]	664418 [0.73%]	564682 [0.64%]	50105 [1.82%]	475902 [3.30%]	7194 [0.64%]	327272 [3.09%]
	N	3240520 [69.47%]	6069579 [55.61%]	1187640 [89.27%]	2461513 [72.69%]	0 [0.00%]	0 [0.00%]	2369470 [86.17%]	6057649 [42.08%]	1023248 [92.18%]	5886927 [55.63%]
	P	177101 [3.79%]	541921 [4.96%]	24606 [1.84%]	114593 [3.38%]	15640265 [17.24%]	14469267 [16.62%]	86980 [3.16%]	891783 [6.19%]	12176 [1.09%]	636877 [6.01%]
	Q	0 [0.00%]	0 [0.00%]	0 [0.00%]	0 [0.00%]	15079950 [16.62%]	15560590 [17.88%]	0 [0.00%]	0 [0.00%]	0 [0.00%]	0 [0.00%]
	T	799503 [17.14%]	2393309 [21.93%]	87730 [6.59%]	531753 [15.70%]	32728077 [36.08%]	32288268 [37.10%]	208540 [7.58%]	4598761 [31.95%]	64211 [5.78%]	2714449 [25.65%]
	X	49301 [1.05%]	137518 [1.26%]	1559 [0.11%]	18217 [0.53%]	0 [0.00%]	0 [0.00%]	3626 [0.13%]	146553 [1.01%]	1524 [0.13%]	91212 [0.86%]

Notes: The first panel presents the mean and standard deviation of the daily returns, as well as the average duration between trades (measured in seconds). The second panel reports the total as well as the number of trades used after data adjustments. The last panel presents the number of (used) trades taking place at the different exchanges: A (AMEX), B (Boston), C (Cincinnati), D (NASD ADF), M (Chicago), N (NYSE), P (Pacific), Q (NASD), T (NASD), X (Philadelphia). The number in brackets denote the corresponding percentages.



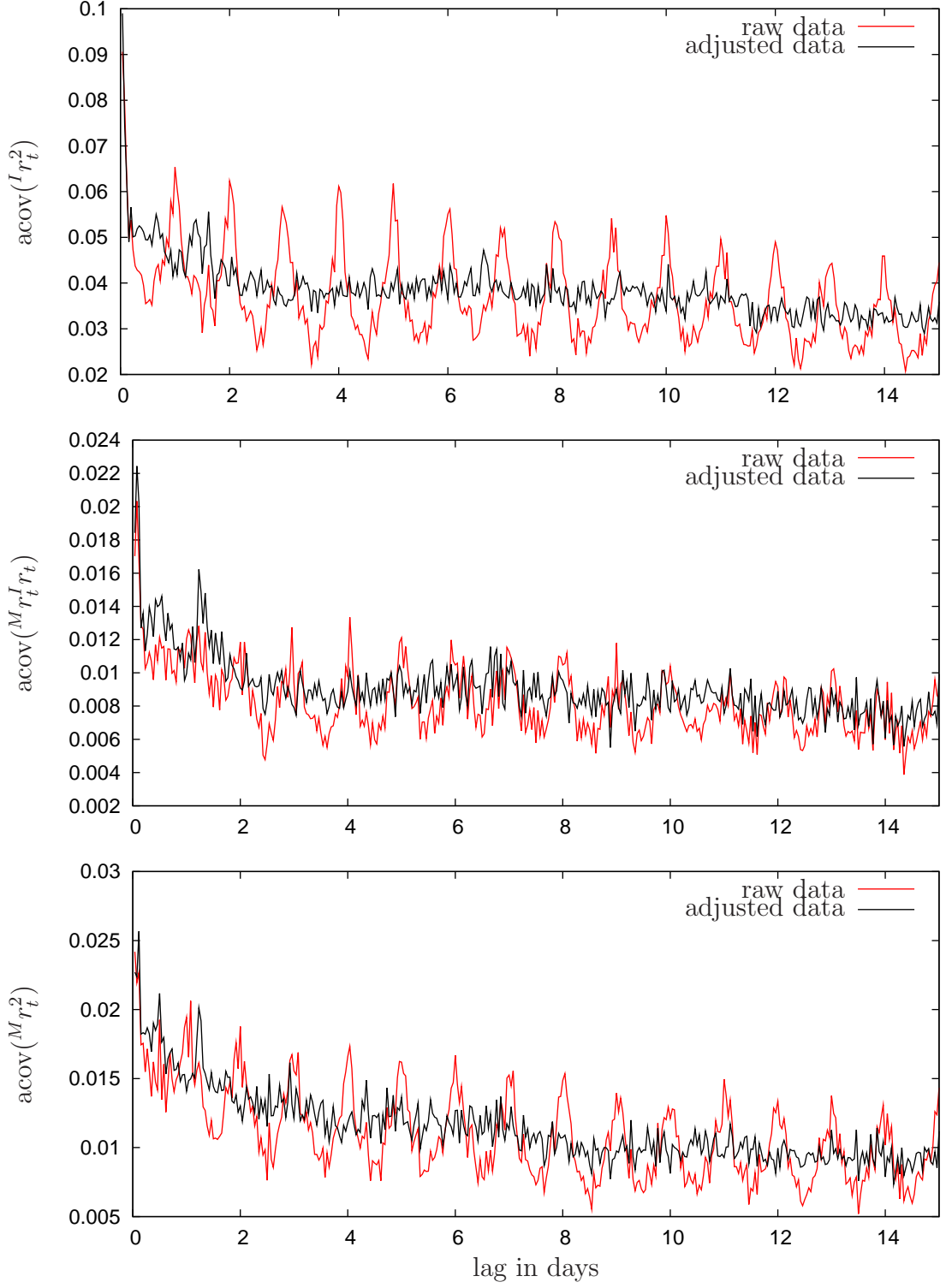
**Figure 2.6:** U-shaped intraday patterns. Depicted is the average number of trades taking place within each minute of the trading day.

adjusted dataset, in order to assess the relevance of cleaning out this intraday pattern a priori.

### Construction of the Realized Covariation Measure

When it comes to the construction of the realized covariation measure we face similar problems as in the univariate case. On one side we may want to use the returns sampled at the highest frequency possible in order to obtain precise estimates, on the other side, market microstructure effects may induce severe biases at very high-frequencies. Moreover, although the literature on how to solve this bias-variance trade-off in the univariate case is by now extensive (see the discussion in the last Section), similarly elaborate approaches for bias-correcting the multivariate measures, i.e. realized covariation, are still pending. This might be due to the just recent introduction of the asymptotic theory for the realized covariation measure. Although some of the methods developed for realized variation can be extended to the multivariate case, the conclusions may be different as market microstructure noise, e.g. noise caused by the bid-and-ask spreads, has different impacts on realized covariation than on realized variance (Voev and Lunde, 2007, see e.g.). In addition, the construction of consistent and efficient realized covariation measures is not only further complicated by the possibility of cross-correlated market microstructure noise, but, importantly, also by the non-synchronous trading of the different assets. In empirical applications the latter problem is usually addressed by applying synchronization methods, such as the last-tick interpolation (e.g. Barndorff-Nielsen and Shephard, 2004a). However, such methods introduce an additional bias to the resulting realized covariation measure, as also discussed in Voev and Lunde (2007). As a consequence, several studies have focused on correcting for this synchronization bias, see e.g. Bandi and Russell (2005b) and Hayashi and Yoshida (2005). A more comprehensive analysis is provided by Voev and Lunde (2007), who derive the asymptotic properties of different synchronization-bias corrected covariation estimators under various forms of market microstructure noise.

However, although some of these estimators are unbiased and consistent under specific noise assumptions, a joint approach for bias-correcting realized variance as well as realized covariation, that is bias-correcting the full covariance matrix, is not yet available in the literature. In fact the estimation of the full covariance matrix is complicated by the different impacts of market microstructure noise on realized variance and covariances. As will become clear in Section 4.4, our estimation methodologies for the multivariate continuous-time stochastic volatility models are based on the full covariance matrix, and we therefore follow Barndorff-Nielsen and Shephard (2004a) and compute the realized covariation estimator according to (2.16) using a lower sampling frequency, whereby we are aware of the potential biases and noisiness of the resulting measure. To reduce these effects, however, we consider 15 minutes or longer time intervals across all assets and apply the last tick interpolation. Moreover, selecting for our empirical analysis the four assets with



**Figure 2.7:** Autocovariance function of 15min squared high-frequency returns of MSFT and INTC and their crossproducts for the raw and adjusted bivariate dataset.

the highest liquidity, i.e. MSFT, INTC, C and PFE,<sup>5</sup> the noise induced bias in the realized-covariance measure should be negligible at this frequency. Moreover, since the durations of the different assets are very similar and small, we expect the synchronization bias to be very small.

### Empirical Properties

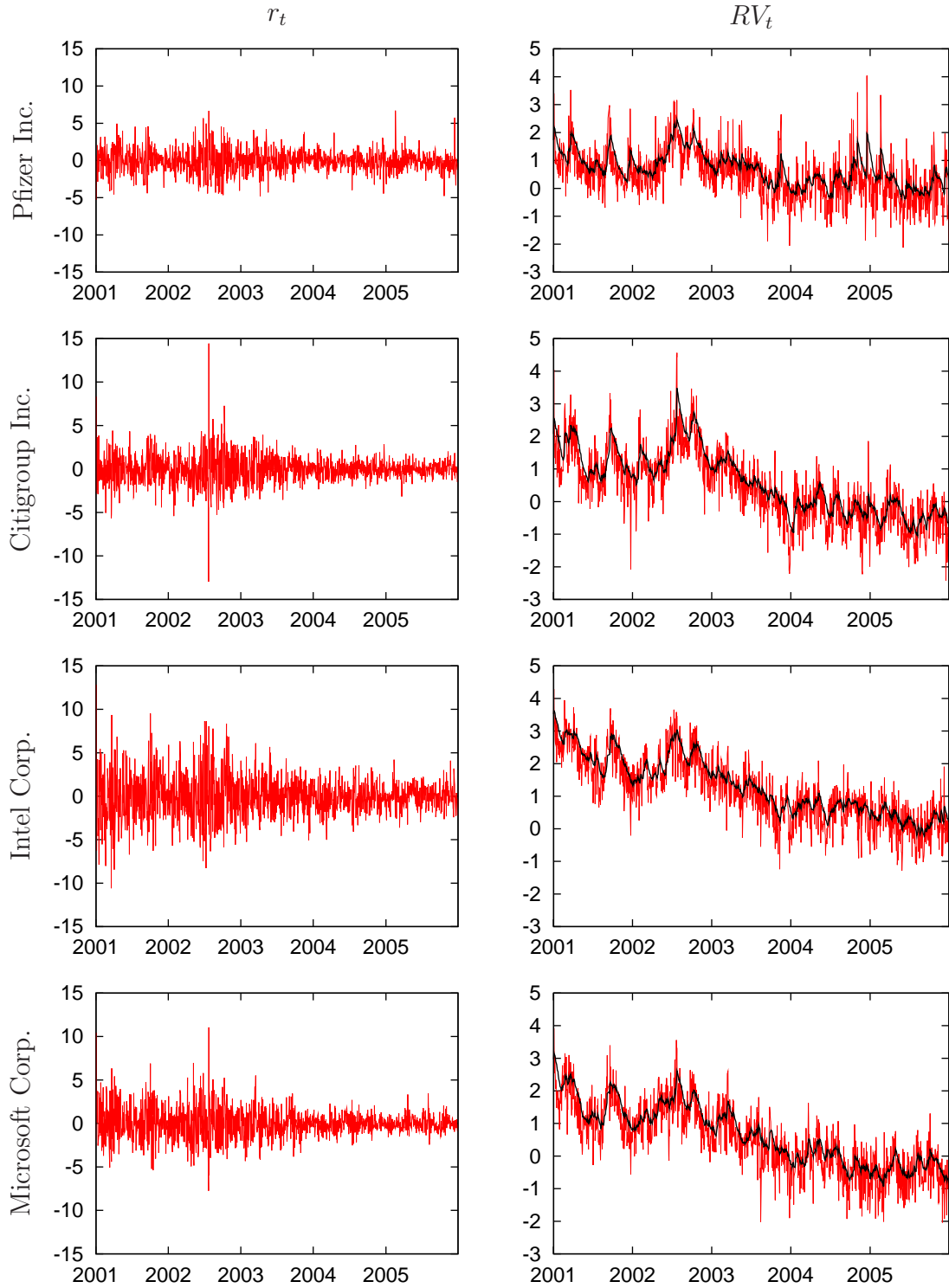
In the following we analyze the empirical properties of the resulting daily return and realized variance and covariance series. Figure 2.8 depicts the the returns (left panel, in percentages) and the corresponding logarithmic realized variances (right panel). All return series show the well-known volatility clustering behavior and seem to move together. In particular, high and low volatility periods seem to occur at the same time across the different assets, whereby we can observe an overall higher volatility during the first half of the sample period when compared to the second half (with the exception of Pfizer Inc.). This becomes even more obvious from the time evolution of the realized variance series. Note that the series are somewhat noisy, as they are based on 15 minutes intervals rather than an infinite sampling frequency.<sup>6</sup> In order to facilitate the visual comparison of these series, we therefore apply an exponential smoother, i.e. we compute the trend realized variance series (displayed in black) by  $\overline{[y_i]_n^{(M)}} = 0.1[y_i]_n^{(M)} + 0.9\overline{[y_i]_{n-1}^{(M)}}$  for the respective asset  $i$ . Figure 2.8 shows the logarithmic realized correlation series, i.e. between asset 1 and 2 we compute  $\frac{\sum_{j=1}^M \mathbf{Y}_{(1)j,n} \mathbf{Y}_{(2)j,n}}{\sqrt{\sum_{j=1}^M \mathbf{Y}_{(1)j,n}^2 \sum_{k=1}^M \mathbf{Y}_{(2)k,n}^2}}$ , along with the correspondingly exponentially smoothed series. The series are nearly throughout non-negative and confirm our previous finding of a positive comovement across the different assets. Moreover, as one might have conjectured from the previous plots, Pfizer Inc. shows slightly less and more volatile correlations with the other stocks, whereas INTC and MSFT exhibit the strongest correlation as expected.

These visual impressions are confirmed by the summary statistics reported in Table 2.4. In particular, the means of the realized correlations of Pfizer Inc. are lower than the other correlations while their variances exceed those of the others, indicating that Pfizer Inc. moves less closely with the other stocks. Overall, the unconditional distributions of the realized correlations are only slightly skewed to the left and exhibit only weak excess kurtosis, i.e. being close to Gaussianity, with the only exception being the realized correlation between INTC and MSFT, for which we observe larger skewness and slight fat tails.

The descriptive statistics for the daily returns of the different assets show the commonly observed skewness and fat-tailedness of the unconditional distribution. Note also, that the skewness and leptokurtosis are less pronounced for single assets

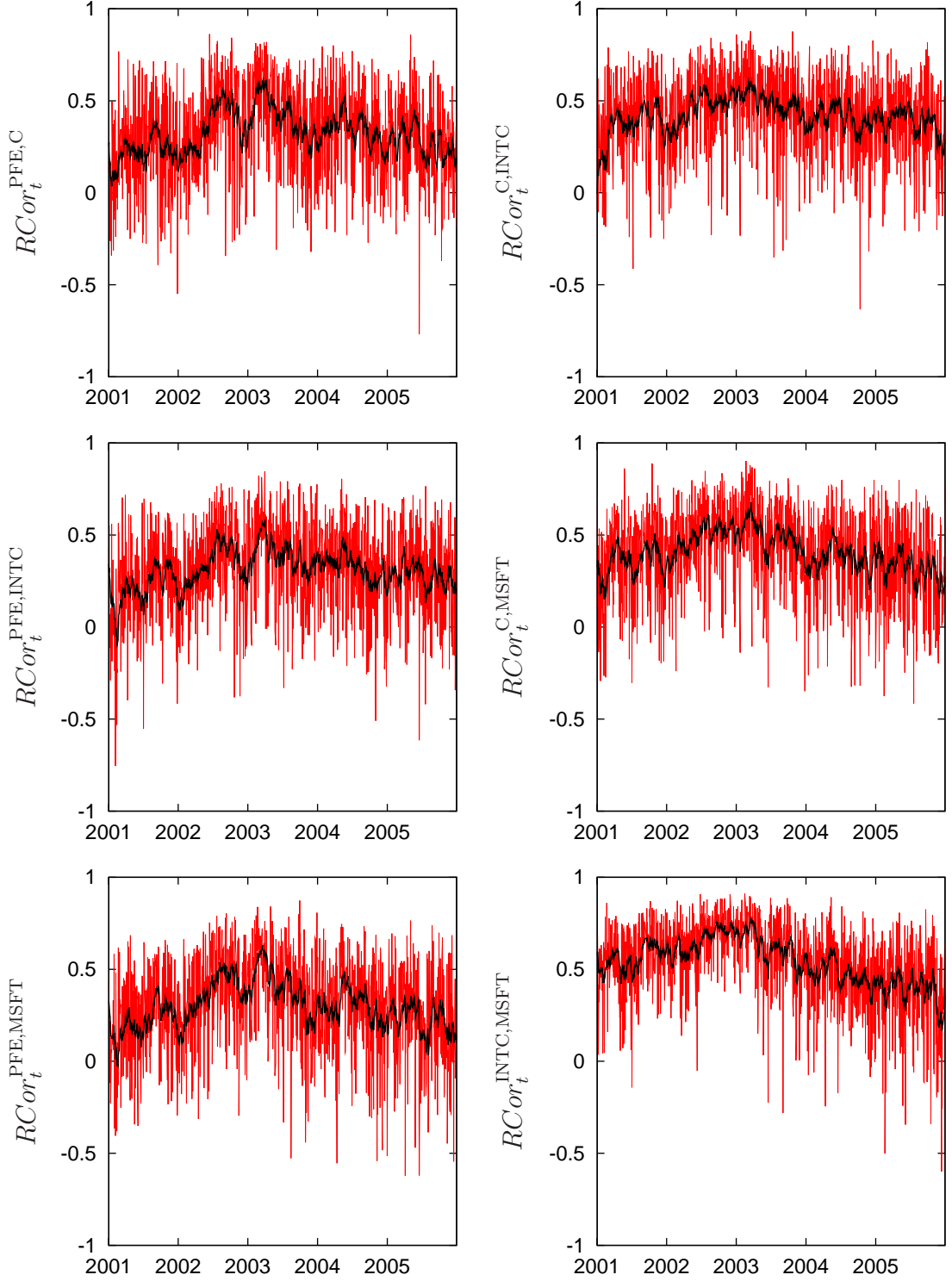
<sup>5</sup>Note that the least traded asset of our sample still exhibits a transaction on average every two and a half seconds.

<sup>6</sup>Such pattern has already been reported in Barndorff-Nielsen and Shephard (2005, 2004a), illustrating via a simulation study that the realized variance and covariance errors, respectively, increase with lower sampling frequencies.



**Figure 2.8:** Daily returns and logarithmic realized variances. The left panel of the figure shows the time evolution of the daily returns of the different assets, whereas the right panel exhibits the corresponding realized variance series, whereby the black line depicts the exponentially smoothed realized variance series (see text for more details on the particular exponential smoother used).





**Figure 2.9:** Daily realized correlations. The figure shows the pairwise realized correlations between the different assets along with the exponentially smoothed realized correlation series (black lines) (see text for more details on the particular exponential smoother used).

**Table 2.4:** Descriptive Statistics of the Multivariate Dataset (C, INTC, MSFT, PFE)

Series	Mean	Std.Dev.	Median	Skewness	Exc.Kurt.	Ljung-Box(10)
$r^{\text{PFE}}$	-0.0979	1.5244	-0.1475	0.1917	1.5139	14.6487
$r^{\text{C}}$	-0.0392	1.6201	-0.0891	0.3253	9.2576	23.4088
$r^{\text{INTC}}$	-0.0462	2.3873	-0.1157	0.2804	2.0375	17.9267
$r^{\text{MSFT}}$	0.0108	1.7308	-0.0603	0.5278	3.4031	18.5505
$RV^{\text{PFE}}$	2.4446	3.4322	1.5344	6.6163	70.2943	709.9982
$RV^{\text{C}}$	2.9525	5.8487	1.4786	9.5613	126.8565	1876.1827
$RV^{\text{INTC}}$	5.4312	6.4778	2.9878	3.1128	15.8839	4687.1599
$RV^{\text{MSFT}}$	2.9007	3.8369	1.5814	4.2629	30.5911	3427.8917
$RCor^{\text{PFE,C}}$	0.3185	0.2533	0.3354	-0.3456	-0.2365	385.4888
$RCor^{\text{PFE,INTC}}$	0.3071	0.2432	0.3299	-0.4826	0.1765	255.5214
$RCor^{\text{PFE,MSFT}}$	0.2987	0.2632	0.3200	-0.5048	-0.0083	400.6457
$RCor^{\text{C,INTC}}$	0.4217	0.2234	0.4521	-0.7204	0.5624	151.6455
$RCor^{\text{C,MSFT}}$	0.4037	0.2379	0.4346	-0.6390	0.0590	265.0768
$RCor^{\text{INTC,MSFT}}$	0.5362	0.2231	0.5770	-1.0801	1.5784	861.5205

than for a stock market index, as is also revealed by comparing the respective values with those of the S&P500 index futures reported in Table 2.1, which however is also based on a different sample period. Moreover, the single assets exhibit less, in fact nearly insignificant serial correlation, as indicated by the Ljung–Box statistics for up to tenth order autocorrelation.<sup>7</sup>

According to Table 2.4 the unconditional distribution of the realized variance series of the single assets considered in this study is less skewed and leptokurtic than the realized variance of the stock market index (see table 2.1), but still significant. In addition, we also find highly significant own serial dependencies as indicated by the corresponding Ljung–Box statistics.<sup>8</sup> Similar results are found for the realized correlations.

---

<sup>7</sup>The corresponding critical value at the five and one percent significance levels is 18.31 and 23.21, respectively.

<sup>8</sup>This is also supported by the sample autocorrelation functions not presented here. Moreover, the sample autocorrelation functions also exhibit a hyperbolic decay.

### 3 Statistical Assessment of Univariate Continuous-Time Stochastic Volatility Models

Modeling the dynamics of asset prices is crucial for derivative pricing and an adequate risk management. As a consequence, a plethora of different continuous-time stochastic volatility models has been developed that aim at capturing the stylized facts of stock returns. Although modeling asset prices by stochastic differential equations was already introduced by Bachelier (1900), this approach just gained further attention through the work of Black and Scholes (1973) and Merton (1973) showing that the continuous-time specification facilitates derivative pricing. However, by now, it is well-known, that the Black-Scholes model, that assumes a geometric Brownian motion for the asset price process, is unable to reproduce the stylized facts of stock returns, in particular the asymmetry and fat tails of the unconditional return distribution, as well as the time-variation and persistence in the volatility with high/low volatile periods following high/low volatile periods, the so-called volatility clustering. As a consequence different extensions have been proposed in the literature. E.g. Merton (1976) already included a jump process into the price diffusion process, which can account for the observed extreme outliers. Heston (1993), Hull and White (1987) and Scott (1987) were the first to introduce continuous-time stochastic volatility models, in which volatility clustering is introduced by specifying an extra random, persistent process for the instantaneous volatility. A combination of both approaches has been examined by e.g. Bates (1996a,b), Bakshi et al. (1997) and Dai and Singleton (2000) as well as Andersen et al. (2002), inter alia. Further extensions are the multi-factor models, in which the volatility is commonly given either by an affine or logarithmic function of these factors (e.g. Andersen et al., 2002; Chernov et al., 2003; Dai and Singleton, 2000); or the additional inclusion of jumps into the volatility specifications such as in Bates (2000), Duffie et al. (2000), Pan (2002) and Eraker et al. (2003), which encounters the fact observed by Jones (2003), that a large shock to volatility of returns leads to a rapid increase in the volatility itself. Whereas all of these models specify either a pure diffusion process or a jump-diffusion process for returns and volatility, whereby the jumps are usually modeled by a Poisson process, the use of other Lévy processes has also become popular over the recent years, see e.g. the non-Gaussian Ornstein-Uhlenbeck-type stochastic volatility models proposed by Barndorff-Nielsen and Shephard (2001b), the continuous-time autoregressive mov-

ing average (CARMA) process as proposed by Brockwell (2001) and also considered in Todorov and Tauchen (2005), or the time–changed Lévy processes of Carr et al. (2003) and Huang and Wu (2004).

Given the large number of different continuous–time stochastic volatility models, it is interesting to assess their ability to reproduce the stylized facts. Generally, most of the existing empirical studies primarily focus on the model assessment and comparison of the affine and logarithmic jump–diffusion models. Using daily or lower–frequency data, the empirical results, however, do not allow for a very clear distinction between pure diffusion multi–factor stochastic volatility models and lower–order models with jumps. In view of the often observed large intraday price–movements, one might conjecture that the daily data might just not be informative enough to provide a distinction between the models. Moreover, as we have seen in the previous chapter, high–frequency data provides indeed new information on the price process and we therefore re–assess the affine and logarithmic jump–diffusion models using realized variation measures.

Hence, our main objective is to evaluate and compare the ability of these models to reproduce the stylized facts of returns and realized variations within a unified framework. In particular, we use the general scientific modeling (GSM) method recently proposed by Gallant and McCulloch (2005). The main idea of GSM is that the usually unavailable transition density of the continuous–time stochastic volatility model is expressed in terms of the transition density of a highly accurate auxiliary model. Importantly, however, in contrast to other indirect inference methods, GSM additionally allows to incorporate some prior information (e.g. on the unconditional mean of the returns). Moreover, the continuous–time models can be interpreted as prior information on the parameter space of the auxiliary model. As such, model assessment is strongly simplified as it can be conducted in terms of the auxiliary model, which provides a unifying framework for model assessment.

The remainder of the chapter is structured as follows. The next section reviews the affine and logarithmic jump–diffusion models and introduces the model specifications we consider in our analysis. Section 3.2 discusses the GSM method. As it requires a highly accurate auxiliary model, we introduce in Section 3.3 a discrete–time model for daily returns, realized variations and jumps, whereby the adequacy of this model is illustrated through a detailed residual analysis and a simulation study. Section 3.4 discusses the prior information we impose, and Section 3.5 provides the estimation and some simulation results. Section 3.6 provides a summary of this chapter.

## 3.1 Model Specifications

In this section we discuss the univariate affine and logarithmic continuous–time stochastic volatility models. The models we consider can be nested in the following

representation, whereby we adopt a similar notation as in Chernov et al. (2003)

$$dy_t = (\alpha_{10} + \alpha_{12}U_{2t})dt + \sigma(\beta_{10} + \beta_{13}U_{3t} + \beta_{14}U_{4t}) \quad (3.1)$$

$$\begin{aligned} & \times \left( \sqrt{1 - \psi_{13}^2 - \psi_{14}^2} dW_{1t} + \psi_{13} dW_{3t} + \psi_{14} dW_{4t} \right) + J_{1,t} dN_t, \\ dU_{2t} &= \alpha_{22}U_{2t}dt + dW_{2t}, \end{aligned} \quad (3.2)$$

with  $y_t$  denoting again the logarithmic price of the financial asset, whose drift is allowed to be driven by a Gaussian Ornstein–Uhlenbeck process, denoted by  $U_{2t}$ , in order to capture potential short–term dynamics in the returns. Moreover, the price follows either a pure diffusion or a jump–diffusion process. The inclusion of jumps into the price process was already proposed by Merton (1976) and should help to capture outliers in the returns. Whereas Merton assumed a constant volatility coefficient, Heston (1993) and Scott (1987) suggested a stochastic volatility specification without jumps. In most empirical applications, however, it turned out that one volatility factor might not be sufficient enough to accommodate for both the extreme outliers as well as the volatility persistence. As a consequence, multi–factor stochastic volatility models or stochastic volatility models with jumps have been proposed. Here, the coefficient  $\sigma(\cdot)$  of the diffusion part is given by a linear combination of at most two mean–reverting pure diffusion volatility factors denoted by  $U_{3t}$  and  $U_{4t}$ . Depending on the functional form of the volatility coefficient we either obtain the affine or the logarithmic stochastic volatility specification, as is described below. The observed asymmetry in the return process or the so–called leverage effect is usually introduced by correlating the Brownian Motions in the price and the volatility specifications. Note, that the parameters  $\psi_{ii}$  are the correlation coefficients due to the reparametrization of  $\psi_{11} = \sqrt{1 - \psi_{13}^2 - \psi_{14}^2}$ . In the presence of multiple volatility factors these correlation coefficients cannot be interpreted as the leverage effect. In particular, Chernov et al. (2003) show, that in this case the leverage effect is dependent on the states of the volatility factors.

### 3.1.1 Affine Models

The affine stochastic volatility models have been extensively discussed in i.e. Andersen et al. (2002), Chernov et al. (2003) and Dai and Singleton (2000). In particular, they are characterized by a linear drift and volatility function. The most popular volatility specification in this class is given by

$$\sigma(u) = \sqrt{u} \quad (3.3)$$

assuming non–negative processes for the volatility factors. Commonly, the volatility factors are driven by the following affine specification

$$dU_{it} = (\alpha_{i0} + \alpha_{ii}U_{it})dt + \sqrt{\beta_{i0} + \beta_{ii}U_{it}}dW_{it}, \quad i = 3, 4, \quad (3.4)$$

i.e. a Square–Root or also called Cox–Ingersoll–Ross process. This specification nests the well–known stochastic volatility model of Heston (1993), in which the drift

is only deterministic, i.e. (3.2) is shut down, and there exists only one Square–Root volatility factor. This model is particularly attractive since closed–form expressions for option pricing formulas exist. Moreover, the additivity of the two volatility factors in this specification allow to interpret one factor as a long–memory component accounting for the persistence and a short–memory component accounting for the extreme events (see e.g. Chernov et al., 2003; Engle and Lee, 1999).

### 3.1.2 Logarithmic Models

In the logarithmic models the volatility coefficient is specified by

$$\sigma(u) = \exp(u) \quad (3.5)$$

which shows the familiarity of this model class with the standard discrete–time stochastic volatility and EGARCH models. The volatility factors are given by Gaussian Ornstein–Uhlenbeck processes, i.e.

$$dU_{it} = (\alpha_{i0} + \alpha_{ii}U_{it})dt + dW_{it}, \quad i = 3, 4. \quad (3.6)$$

Obviously, in this specification the volatility factors enter the diffusion component of the price process in a multiplicative way, whereby the exponential function allows for very high volatility values. Also, this specification nests one of the classical stochastic volatility models. In particular, excluding the stochastic drift specification and considering only a one factor–diffusion process we obtain the model of Scott (1987). Note, however, that the log–linear models are under the viewpoint of derivative pricing less attractive than the affine specification, as it requires numerical computations to derive option prices.

### 3.1.3 Jump–Diffusion Models

An alternative way to simultaneously capture the tail thickness and the persistence in volatility is given by additionally allowing for jumps in the return process. Generally, in the jump–diffusion stochastic volatility models the volatility of returns is then restricted to be driven by only one factor ( $U_{3t}$ ), which is responsible for the persistence in volatility whereas the jumps account for the fat tail behavior. Following the specifications considered in Andersen et al. (2002) and Chernov et al. (2003), we specify the jumps in the return process by a Poisson jump process with constant jump intensity, i.e.

$$N_t \sim Poi(\lambda_J) \text{ and jumps sizes } J_{1,t} \sim N(\mu_J, \sigma_J^2). \quad (3.7)$$

The assumption of constant jump intensity can be relaxed, such as in Andersen et al. (2002) and Chernov et al. (2003), who specify the jump intensity as a function of the spot volatility or of the size of the previous jump. In this study, however, we will not consider a time–varying jump intensity as we want to assess the relevance and adequacy of one of the simplest jump specifications versus pure diffusion models.

### 3.1.4 Model Definitions

In this section we present our model abbreviations and the normalizations imposed for identification. Again, for the ease of comparison the abbreviations are adopted from Chernov et al. (2003). In particular for all models we set

$$\alpha_{20} = 0 \text{ and } \beta_{20} = 1. \quad (3.8)$$

For the affine models we impose the following restrictions

$$\beta_{10} = \beta_{30} = \beta_{40} = 0 \text{ and } \beta_{33} = \beta_{44} = 1. \quad (3.9)$$

For the logarithmic models the following constraints apply

$$\alpha_{30} = \alpha_{40} = 0 \text{ and } \beta_{30} = \beta_{40} = 1. \quad (3.10)$$

Moreover, we impose the standard conditions to ensure existence and non-explosivity. Our models are summarized in Table 3.1.

## 3.2 Estimation Methodology

The estimation of continuous-time stochastic volatility models is well-known to be nontrivial due to the presence of unobservable state variables as well as the nonexistence of closed-form expressions for the transition densities. Moreover, the models can only be evaluated with discretely observed data introducing the problem of discretization bias. As a consequence, considerable research effort has been conducted to develop adequate estimation methods ranging over simulated maximum likelihood methods (see e.g. Durham and Gallant, 2002; Pedersen, 1995), quasi maximum likelihood estimation using high-frequency information (e.g. Barndorff-Nielsen and Shephard, 2002a, 2006b), MCMC methods (e.g. Eraker, 2001; Eraker et al., 2003), to method of moments based approaches (e.g. Bibby and Sørensen, 2001; Bollerslev and Zhou, 2002; Duffie and Singleton, 1993; Gallant and Tauchen, 1996; Hansen and Scheinkman, 1995), as well as indirect inference methods (e.g. Gallant and Tauchen, 1996; Gouriéroux et al., 1993; Gallant and McCulloch, 2005).

The use of high-frequency financial data to evaluate the continuous-time stochastic volatility models, however, further increases the complexity in the implementation of some of those methods. For example, the simulated maximum likelihood estimation or MCMC methods can only be applied directly to the high-frequency return data itself (ideally adjusted for market microstructure and diurnal effects), and will therefore in general involve irregularly spaced time-intervals. Instead, method of moments and indirect inference methods can exploit the high-frequency information contained in regularly sampled variation measures. However, their feasibility relies on the availability of sufficient moment conditions for those measures, or on the existence of a discrete-time parametric auxiliary model, that adequately describes the dynamics and distributional characteristics of these series. Fortunately, such models are available, and the model proposed in Bollerslev et al. (2007)



**Table 3.1:** Jump–Diffusion Model Specifications

Abb.	Model description	$\alpha_{10}$	$\alpha_{12}$	$\alpha_{22}$	$\alpha_{30}$	$\alpha_{33}$	$\alpha_{40}$	$\alpha_{44}$	$\beta_{10}$	$\beta_{13}$	$\beta_{14}$	$\beta_{30}$	$\beta_{33}$	$\beta_{40}$	$\beta_{44}$	$\psi_{13}$	$\psi_{14}$	$\lambda_J$	$\mu_J$	$\sigma_J$
AFF1V	affine, 1 vola. factor	*	*	*	*	*				*			1			*				
AFF2V	affine, 2 vola. factors	*	*	*	*	*	*	*		*	*		1		1	*	*			
AFF1V–J	affine, 1 vola. factor plus jumps	*	*	*	*	*				*			1			*		*	*	*
LL1V	log., 1 vola. factor	*	*	*		*			*	*		1				*				

$$\begin{aligned}
dy_t &= (\alpha_{10} + \alpha_{12}U_{2t})dt + \sigma (\beta_{10} + \beta_{13}U_{3t} + \beta_{14}U_{4t}) \\
&\quad \times \left( \sqrt{1 - \psi_{13}^2 - \psi_{14}^2}dW_{1t} + \psi_{13}dW_{3t} + \psi_{14}dW_{4t} \right) + J_{1,t}dN_t, \\
dU_{2t} &= \alpha_{22}U_{2t}dt + dW_{2t}, \\
dU_{it} &= (\alpha_{i0} + \alpha_{ii}U_{it})dt + (\beta_{i0} + \beta_{ii}U_{it})^{\gamma_i}dW_{it}, \quad i = 3, 4 \\
\sigma(u) &= \sqrt{u}, \quad \gamma_i = 0.5 \text{ for } \mathbf{affine} \\
\sigma(u) &= \exp(u), \quad \gamma_i = 0 \text{ for } \mathbf{logarithmic} \\
N_t &\sim Poi(\lambda_J), \\
J_{1,t} &\sim N(\mu_J, \sigma_J^2)
\end{aligned}$$

provides an extensive and highly accurate description of the dynamics and inter-relationships of returns, and the continuous–time and jump components of total price variation.

Hence, we find ourselves in a situation to which indirect inference methods are particularly well–suited, i.e. the transition density of the models under consideration are not available in closed–form, however, one can simulate from those models. Moreover, there exists a parametric auxiliary model that provides an adequate description of the data. Candidate estimation methods therefore include the efficient method of moments of Gallant and Tauchen (1996), or the general scientific modeling (GSM) method proposed in Gallant and McCulloch (2005). In this work we use the GSM method, since it allows us to additionally incorporate prior information, and, more importantly, provides tools to assess the ability of the candidate continuous–time stochastic volatility models to capture the stylized facts of the high–frequency based data. In contrast to the EMM method,<sup>1</sup> the interesting data features must not be associated with specific parameters of the auxiliary model, but can also be given by functionals that cannot be expressed in terms of the auxiliary model’s parameters, such as e.g. the mean or variance of the observed series. We continue the discussion of our estimation approach by describing the GSM method in Subsection 3.2.1, and by introducing the highly accurate auxiliary model of Bollerslev et al. (2007) for the high–frequency information contained in the realized variation measures in Subsection 3.3.1.

### 3.2.1 The General Scientific Modeling Method

Let us first briefly sketch the main idea of the general scientific modeling method of Gallant and McCulloch (2005), which is especially suited for the estimation of models with unknown transition density. In particular, GSM assumes that in addition to the structural model, which is the model of interest, there also exists a parametric auxiliary or statistical model that provides a highly accurate description of the data and for which the transition density is available in closed–form. Then, these two models are linked to each other by the existence of a map, mapping the parameters of the structural model into the parameter space of the auxiliary model. This allows to express the unknown transition density of the structural model in terms of the transition density of the auxiliary model. Hence, given the auxiliary model and the map, the parameters of the structural model can be estimated by simple maximum likelihood methods. However, the map is usually unknown and needs to be computed. In GSM it is therefore additionally assumed that one can simulate from the structural model. The map is then computed by fitting the auxiliary model to the simulated data. The assumptions imposed so far, i.e. the existence of a highly accurate auxiliary model with known form of the transition density, the existence of the map, as well as the possibility to simulate from the structural model, show

---

<sup>1</sup>In EMM the score t–ratios can be used to assess how well the structural model fits the stylized facts of the data.

the familiarity of the GSM method to other indirect inference approaches. However, GSM additionally allows to incorporate also prior information, such that the method can be interpreted as a Bayesian indirect inference approach. In this case, the main objective is to compute the posterior of the parameters of the structural model based on the auxiliary model and the map, which involves standard MCMC methods. In the following we will discuss GSM in more detail and introduce the notation, whereby we follow Gallant and McCulloch (2005).

Let the transition density of the structural model, i.e. of the continuous–time stochastic volatility model, be denoted by

$$p(y_t|x_{t-1}, \theta) \quad \theta \in \Theta,$$

whereby  $x_{t-1}$  subsumes all relevant past information of the process  $y_t$ . As previously noted, the form of the transition density is unknown, but one can simulate from the structural model for given  $\theta$ . Moreover, there might exist some prior information on the structural model that can be expressed either in terms of  $\theta$  or by simulating from the model. The latter functionals or characteristics of the process, such as for example moments of the unobserved drift or volatility factors, are denoted by  $\Psi : p(\cdot|\cdot, \theta) \mapsto \psi$  and we can summarize the prior information on the structural model by  $\pi(\theta, \psi)$ .

In contrast, the transition density of the parametric auxiliary model is known in closed–form and is denoted by

$$f(y_t|x_{t-1}, \eta) \quad \eta \in H.$$

Prior information on the parameters or other functionals of the auxiliary model  $\Upsilon : f(\cdot|\cdot, \eta) \mapsto v$  can be imposed as well and are expressed by  $\pi(\eta, v)$ . In general, the functionals  $v$  can be computed by simulations from the auxiliary model, but can also include elements of  $\psi$ , such as moments of observed variables. This is very appealing, since very often one has some prior information about unconditional moments of the return series etc.

Moreover, it is assumed, that the auxiliary model provides an accurate description of the process  $\{y_t\}$ , and that there exists a map  $g : \theta \mapsto \eta$ , i.e. mapping the parameters of the structural model into the parameter space of the auxiliary model. Given this map, the transition density of the structural model can be computed using the transition density of the auxiliary model

$$p(y_t|x_{t-1}, \theta) = f(y_t|x_{t-1}, g(\theta)) \quad \theta \in \Theta. \quad (3.11)$$

Hence, the unknown likelihood of the structural model can be inferred from the likelihood of the auxiliary model with parameters restricted through the map. Since the auxiliary model is generally highly parametrized its dimension of  $\theta$  is larger than that of the structural model, and the map therefore generates a manifold  $\mathcal{M}$  on  $\eta$ . So, given the map the analysis can be carried out in the parameter space of the auxiliary model.

In general, however, the map is unknown and needs to be computed by minimizing the Kullback–Leibler divergence between  $p(y|x, \theta)$  and  $f(y|x, \eta)$ . This involves the computation of  $\int \int \log f(y, \eta) p(y|x, \theta) dy dp(x|\theta) dx$ , which can be approximated by  $\frac{1}{N} \sum_{t=1}^N \log f(\hat{y}_t|\hat{x}_{t-1}, \eta)$  with  $\{\hat{y}_t, \hat{x}_{t-1}\}$  denoting a sequence of data of length  $N$  simulated according to  $p(y_t|x_{t-1}, \theta)$ . The minimization problem then results into the maximum likelihood estimation of the auxiliary model on the simulated data and the map is therefore given by

$$g : \theta \mapsto \arg \max_{\eta} \sum_{t=1}^N \log f(\hat{y}_t|\hat{x}_{t-1}, \eta). \quad (3.12)$$

Therefore, in applications the map can be uncovered by simulating for any given  $\theta$  from the structural model, and by fitting the auxiliary model to the simulated data. The resulting maximum likelihood estimator  $\eta$  then corresponds to the parameter vector  $\theta$  given through the map.

Once the map is known, i.e. once we have computed the maximum likelihood estimator  $\eta$ , we can compute for each  $\theta$  the likelihood of the structural model at the observed data  $(\tilde{y}_t)$  by

$$\mathcal{L}(g(\theta)) = \prod_{t=1}^T f(\tilde{y}_t|\tilde{x}_{t-1}, g(\theta)). \quad (3.13)$$

Then the remaining steps to obtain the parameter estimates of the structural model are straightforward.

In particular, additionally allowing for prior information (subsumed in  $\pi(\theta, \psi, \eta, \nu)$ ) yields a Bayesian estimation approach and the posterior of the parameters of the structural model  $p(\theta|y, x) \propto \mathcal{L}(g(\theta)) \pi(\theta, \psi, \eta, \nu)$  can be computed via standard MCMC methods. More specifically, a Metropolis Hastings algorithm can be employed that only differs from the standard algorithms by the additional steps involved to compute the map, such that the likelihood of the structural model can be approximated by  $\mathcal{L}(g(\theta))$ . The algorithm suggested in Gallant and McCulloch (2005) is as follows

1. Draw  $\theta^*$  according to  $q(\theta^*|\theta^o)$
2. Draw  $\{\hat{y}_t, \hat{x}_{t-1}\}_{t=1}^N$  according to  $p(y_t|x_{t-1}, \theta^*)$ .
3. Compute  $\eta^* = g(\theta^*)$  and  $\psi^*$  from the simulation  $\{\hat{y}_t, \hat{x}_{t-1}\}_{t=1}^N$  and  $\nu^*$  from  $\eta^*$ .
4. Let  $\alpha = \min \left( 1, \frac{\mathcal{L}(g(\theta^*))\pi(\theta^*, \psi^*, \eta^*, \nu^*)q(\theta^*|\theta^o)}{\mathcal{L}(g(\theta^o))\pi(\theta^o, \psi^o, \eta^o, \nu^o)q(\theta^o|\theta^T)} \right)$
5. Accept  $(\theta' = \theta^*)$  with probability  $\alpha$  otherwise repeat  $(\theta' = \theta^o)$ .

More details on the algorithm and a possible specification of the proposal density are discussed in Gallant and McCulloch (2005).

The Bayesian character of the GSM method allows for a particularly detailed model assessment. So far, the estimation procedure build on the assumption that the structural model holds exactly. As a consequence, the parameter estimates of the structural model, or more generally, the whole posterior expressed in the  $\eta$ -space, is restricted to the manifold. This assumption may be relaxed, in order to assess the adequacy of the structural model.

When allowing the structural model to move away from the manifold, its posterior can move in the  $\eta$ -space in search of the likelihood evaluated at the data. If the model is correct, the shape of the posterior may change, however, the mode should still lie on the manifold. In contrast, if the structural model is incorrect, then the location of the posterior will shift and its mode will be off the manifold. Obviously, inference off the manifold can only be carried out in the  $\eta$ -space, since the structural model is not valid anymore.

The implementation of the model assessment procedure is best understood when the structural model is interpreted as imposing a prior on  $\eta$  through the map. Hence, the prior on the statistical model now consists of two components: the first is the prior that  $\eta$  is restricted to the manifold; the second is the prior  $\pi(\eta, v)$ . Now, relaxing the first prior allows the posterior to move away from the manifold. This relaxation can be expressed by a parameter  $\kappa$  and the relaxed prior is then given by

$$\pi_{\kappa}(\eta, v) \propto \pi(\theta_j, \psi_j) \exp \left\{ -\frac{d(\eta, \mathcal{M})}{2\kappa} \right\} \pi(\eta, v) \quad (3.14)$$

with  $d(\eta, \mathcal{M})$  measuring the minimum distance of  $\eta$  from the manifold. A large value of  $\kappa$  then corresponds to a relaxation of the prior.  $\pi(\theta_j, \psi_j)$  denotes the prior information of the structural model on the point on the manifold, for which the distance is minimal. In particular, the distance can be computed by  $d(\eta, \mathcal{M}) = \min_j [\eta - g(\theta_j)]' \Sigma_{\eta}^{-1} [\eta - g(\theta_j)]$  with scaling matrix  $\Sigma_{\eta}$  that ideally puts  $\eta$  on the scaling of the posterior. A proposal for the computation of  $\Sigma_{\eta}$  is given in Gallant and McCulloch (2005).<sup>2</sup>

The inference off the manifold can also be conducted for the parameters of the auxiliary model and therefore provides similar strategies as EMM to assess the ability of the structural model to reproduce the stylized facts of the data associated with those parameters. However, a comparison of the posteriors of other functionals of the auxiliary model is also possible and therefore allows to assess additional data characteristics that cannot be expressed in terms of the parameters of the auxiliary model.

---

<sup>2</sup>Applying an MCMC algorithm to compute the map, the weighting matrix  $\Sigma_{\eta}$  can be computed from the parameter chains obtained for each maximum likelihood estimation of  $\eta$  on the simulated data.

### 3.3 The Auxiliary Model

Recall that our objective is to estimate different continuous–time stochastic volatility models using high–frequency information. Therefore, if we want to apply the GSM method we first need to specify a highly accurate model for high–frequency data. However, as the modeling of high–frequency returns is complicated by market microstructure effects, diurnal volatility patterns etc. (see also our discussion in Section 2.2) we explore the high–frequency information by modeling realized variation measures. In fact, the nonparametric volatility measures have already inspired the development of a series of new and simple–to–implement reduced form volatility forecasting models in which the realized volatilities are modeled by standard discrete–time time series procedures, examples of which include Andersen et al. (2003a, 2007), Corsi (2004), Corsi et al. (2007), Deo et al. (2006), Koopman et al. (2005) and Martens et al. (2004), among others. By effectively incorporating the high–frequency data into the volatility measurements, these simple discrete–time models generally out–perform existing more complicated parametric volatility models based on the corresponding daily return observations only. The simplicity of these methods, however, comes at the cost of disregarding information about the different volatility components. With the exception of Andersen et al. (2007), who simply included lagged measures for the jump component into a univariate linear forecasting model for the total realized variation, none of the above listed studies have made use of the decomposition of the total variation into its separate continuous and jump components. However, as illustrated in Section 2.2 jumps are apparently relevant and there exist distinctly different distributional features and time–series patterns of the continuous volatility component and the jump components. Hence, a more structured approach to realized volatility modeling may be preferable, which in turn will allow us to assess the ability of the different univariate continuous–time stochastic volatility models to reproduce the stylized facts implied by the high–frequency information.

In Bollerslev et al. (2007) we develop such an empirically highly accurate multivariate discrete–time volatility model for the returns and the realized continuous sample path and jump variation measures. Note also, that our joint modeling of the returns and the two volatility components allows us to directly assess the importance of the often documented asymmetric relationship between returns and volatility, and whether the observed leverage effect is caused by a negative correlation of the lagged returns with the current continuous volatility component and/or current jumps.<sup>3</sup> We initially estimate the model equation–by–equation under the implicit assumption that the disturbances are independent across the three equations. However, our univariate estimation results reveal important nonlinear contemporaneous dependencies in the disturbances, which we account in a general recursive simultaneous equation system, explicitly allowing for contemporaneous

---

<sup>3</sup>The recent empirical analysis in Bollerslev et al. (2006b) also points toward the existence of a contemporaneous leverage effect in the form of cross–correlated high–frequency returns and absolute returns.

nonlinear inter–dependencies. The recursive structure also makes simulations from the model easy to implement, which is used for checking different aspects of our final preferred specification. It turns out, that our model indeed provides a highly accurate description of the data. Importantly, despite the general and very flexible structure of the model, its transition density is available. Hence, the model can be used for the estimation of univariate continuous–time stochastic volatility models via indirect inference methods, such as GSM.

In the following section we introduce the formulation of the three basic model equations for the returns, Bipower variation and relative jump series. The resulting equation–by–equation estimates are presented in Section 3.3.2, along with an assessment of the cross–equation dependencies in the disturbances. This illustrates the relevance of contemporaneous dependencies and explains why our final model specification is given by a trivariate simultaneous equation model, which is described in Section 3.3.3 along with a presentation of the corresponding maximum likelihood estimates. Simulations from the model are given in Section 3.3.4 showing the adequacy of the model’s fit.

### 3.3.1 A Discrete–Time Model for Daily Returns and Realized Variations

A burgeon literature dating back to Bollerslev (1987) and French et al. (1987) has been concerned with the modeling of daily speculative returns using GARCH and related stochastic volatility models; see, e.g., the review in Bollerslev et al. (1994). More recently, several studies, including Andersen et al. (2003a), Martens et al. (2004), Martens and Zein (2004), Pong et al. (2004), and Thomakos and Wang (2003) among others, have advocated the use of ARFIMA type models, along with approximate long–memory component type structures in Andersen et al. (2007) and Corsi (2004), for modeling the dynamic dependencies in realized volatilities. However, these same ideas have not been applied yet to the Bipower variation, nor to the relative jump measure considered here.<sup>4</sup> More importantly, we are not aware of any other attempts at jointly modeling the daily  $r_t$ ,  $BV_t$  and  $J_t$  series within a coherent multivariate framework. We first consider the specification for the integrated volatility process as measured by daily Bipower variation, followed by a discussion of our models for the relative jump component and the daily returns, respectively.

#### The Bipower Variation Equation

The realized variation only differs from the Bipower variation (by more than measurement errors) in the presence of jumps. Hence, guided by the recent empirical literature pertaining to the modeling of  $RV_t$  cited above, we will here rely on the

---

<sup>4</sup>In a related context, Andersen et al. (2006) have recently explored the use of ACD type models for characterizing the times between significant, according to the ratio statistic in Huang and Tauchen (2005), jumps.



Heterogeneous Autoregressive, or HAR–RV, type model, originally proposed by Corsi (2004) and successfully employed in a closely related context by Andersen et al. (2007), for describing the dynamic dependencies in the  $BV_t$  series.<sup>5</sup> However, in contrast to the HAR–RV model estimates reported in Corsi (2004) and Andersen et al. (2007), which are based on simple least squares, we shall here rely on more efficient maximum likelihood estimation techniques explicitly accounting for the time-dependent conditional heteroskedasticity in the residuals from the  $BV_t$  model through a separate GARCH type specification for the volatility-of-volatility. A similar estimation approach has also recently been implemented by Corsi et al. (2007).

More specifically, to set up the model we define the logarithmic multiperiod Bipower variation measures by the sum of the corresponding daily logarithmic measures

$$(\log BV)_{t+1-k:t} = \frac{1}{k} \sum_{j=1}^k \log BV_{t-j}, \quad (3.15)$$

where  $k = 5$  and  $k = 22$  correspond to (approximately) one week and one month, respectively.<sup>6</sup> Our HAR–GARCH–BV model then takes the form

$$\begin{aligned} \log BV_t &= \alpha_0 + \alpha_d \log BV_{t-1} + \alpha_w (\log BV)_{t-5:t-1} + \alpha_m (\log BV)_{t-22:t-1} \\ &\quad + \theta_1 \frac{|r_{t-1}|}{\sqrt{RV_{t-1}}} + \theta_2 I[r_{t-1} < 0] + \theta_3 \frac{|r_{t-1}|}{\sqrt{RV_{t-1}}} I[r_{t-1} < 0] + \sqrt{h_t} u_t \\ h_t &= \omega + \sum_{j=1}^q \alpha_j (\log BV_{t-1} - x'_{BV} \beta_{BV})^2 + \sum_{j=1}^p \beta_j h_{t-j} + \sum_{j=1}^s \lambda_j BV_{t-j} \end{aligned} \quad (3.16)$$

The lagged daily, weekly and monthly realized variation measures on the right-hand-side of the  $\log BV_t$  equation could, of course, be augmented with additional terms to account for the possibility of even longer-run dependencies. However, the combination of relatively few volatility components often provide a remarkably close approximation to true long-memory dependencies. The remaining, new vis-à-vis the original HAR–RV model in Corsi (2004), terms explicitly allow for a leverage effect in the volatility through the inclusion of the lagged signed returns. The model also permits a level effect in the GARCH model for the volatility-of-volatility. Lastly, to account for deviations from conditional normality, we allow the errors to follow a normal-mixture distribution

$$u_t \stackrel{iid}{\sim} \begin{cases} N_1(0, 1) & \text{with probability } (1 - p_{u,2}) \\ N_2(\mu_{u,2}, \sigma_{u,2}^2) & \text{with probability } p_{u,2} \end{cases}. \quad (3.18)$$

<sup>5</sup>The HAR model may be seen as an extension of the heterogeneous ARCH model first suggested by Müller et al. (1997).

<sup>6</sup>We follow Corsi (2004) in defining the multi-period logarithmic volatility by the sum of the corresponding one-period logarithmic measures. Almost identical empirical results obtain by using the logarithm of the multi-period realized variances in place of the sum of the logarithms.



Having defined the model for the continuous volatility component, we next turn our attention to the specification of the jump component.

### The Jump Equation

Consistent with the results in Andersen et al. (2007) pertaining to the time series of significant squared jumps, the descriptive statistics in Section 2.2 point toward fairly weak, albeit not zero, own serial dependencies in the relative jump series. To best accommodate this we specify a standard autoregressive model augmented with the same leverage type terms used in the  $BV_t$  equation

$$\begin{aligned} \log \left( \frac{RV_t}{BV_t} \right) &= \delta_0 + \psi_1 \frac{|r_{t-1}|}{\sqrt{RV_{t-1}}} + \psi_2 I[r_{t-1} < 0] + \psi_3 \frac{|r_{t-1}|}{\sqrt{RV_{t-1}}} I[r_{t-1} < 0] \\ &\quad + \sum_{j=1}^n \delta_j \log \left( \frac{RV_{t-j}}{BV_{t-j}} \right) + \nu_t. \end{aligned} \quad (3.19)$$

This in turn allows us to disentangle whether the well-documented asymmetric negative relationship between total volatility and return innovations is primarily driven by the response of the continuous volatility component and/or the reaction of the jump component.

Experimentation suggests that the innovations in the jump equation are well described by a mixture of a zero mean Normal Inverse Gaussian (NIG) distribution and an Inverse Gaussian (IG) distribution

$$\nu_t \stackrel{iid}{\sim} \begin{cases} \text{NIG}_0(\alpha_{NIG}, \beta_{NIG}, \delta_{NIG}) & \text{with probability } (1 - p_{\nu,2}) \\ \text{IG}(\lambda_{IG}, \mu_{IG}) & \text{with probability } p_{\nu,2} \end{cases}. \quad (3.20)$$

Other asymmetric distributions could, of course, be considered, but a mixture based on one distribution having support on the whole real line and the other being defined only on the positive domain appears a natural choice for characterizing the jump innovation distribution. Intuitively, the NIG distribution may be seen as primarily accounting for the small day-to-day fluctuations in the logarithmic realized variance around the logarithmic Bipower variation attributable to measurement errors and small jumps, while the positive IG distribution captures the innovations associated with large genuine jumps, or the right tail of the distribution. Moreover, the NIG and IG distributions both have very flexible shapes, and the superior fit afforded by this particular mixture of distributions is indeed confirmed by our model estimates discussed below.

### The Return Equation

Our final model for the distribution of the daily returns rely on the nonparametric  $RV_t$  measure for capturing the total price variability. This same idea has previously

been used in the context of modeling daily returns by Forsberg and Bollerslev (2002). Note that even though we do not directly model  $RV_t$ , the conditional distribution of the total price variation is readily inferred from our models for the logarithmic Bipower variation and the jumps discussed above based upon the definition in equation (2.13); i.e.,  $RV_t := \exp(J_t + \log BV_t)$ .

Specifically, allowing for up to  $d$ 'th order serial correlation, we postulate the following simple autoregressive model for the daily return process

$$r_t = \gamma_0 + \sum_{j=1}^d \gamma_j r_{t-j} + \sqrt{RV_t} \epsilon_t. \quad (3.21)$$

Our final preferred model takes the innovations to be standard normally distributed

$$\epsilon_t \stackrel{iid}{\sim} \mathcal{N}(0, 1), \quad (3.22)$$

but as discussed further below, we also experimented with other more flexible mixtures-of-distributions to allow for deviations from conditional normality. However, broadly consistent with the summary statistics in Table 2.1, we found that the standard normal distribution provided as good a fit as any of these other distributions.

We next turn to a discussion of the univariate estimation results for this very simple return generating process along with the other two equations for the realized variation measures making up our complete system.

### 3.3.2 Equation-by–Equation Estimation

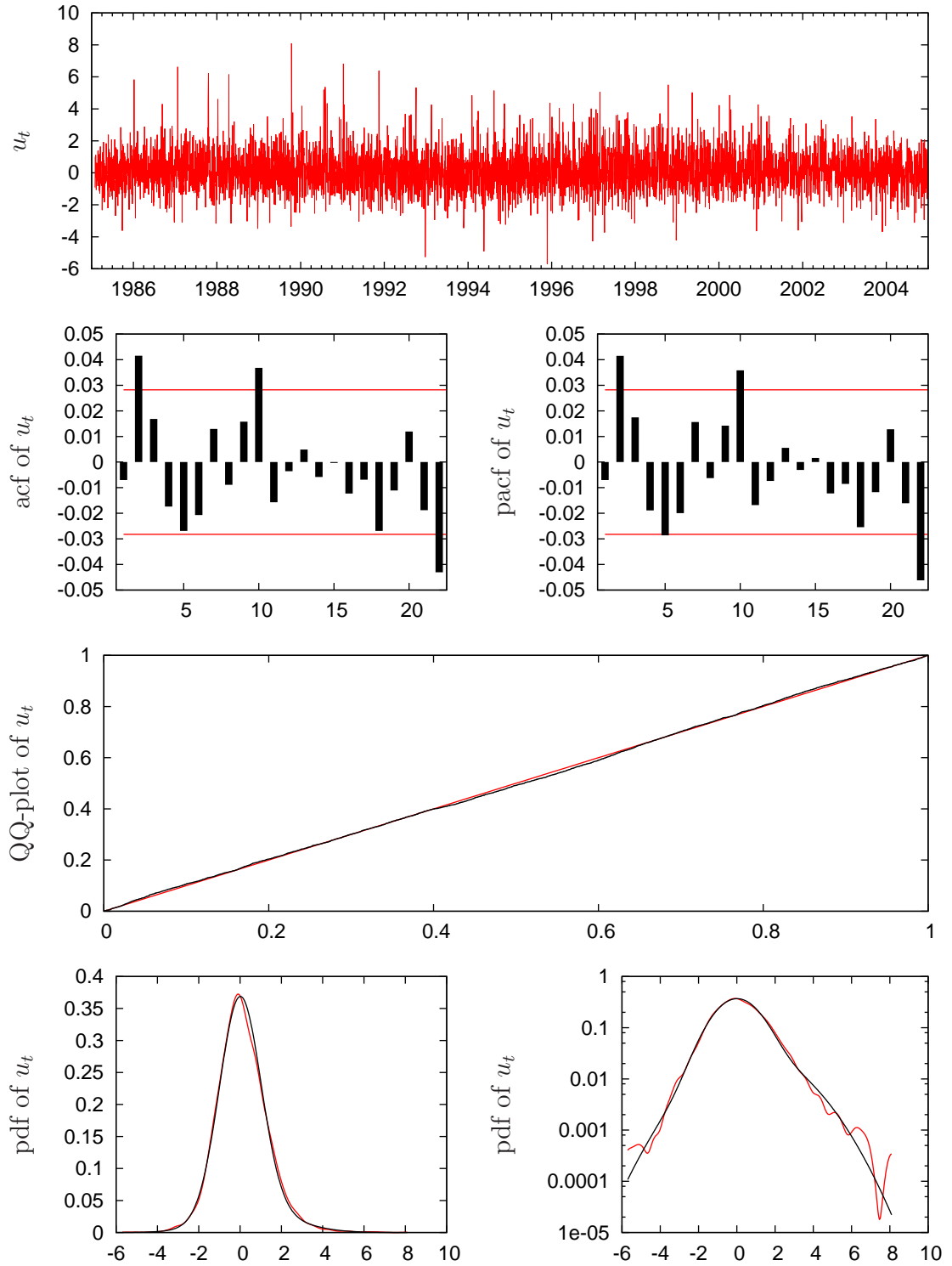
The recursive structure of the three equation system defined in the preceding sections, means that as long as the disturbances are independent across equations, each of the three models may be estimated efficiently in isolation using standard maximum likelihood methods. The validity of the assumption of independent disturbances is, of course, questionable, and we will subsequently investigate the adequacy of this based upon the single-equation estimates.

The estimation results for each of the three equations, along with the corresponding asymptotic standard errors for the parameter estimates, are reported in Table 3.2. Figures 3.1 to 3.3 show the resulting residuals, their autocorrelation and partial autocorrelation functions, as well as the QQ plots and kernel density estimates. The selection of the autoregressive lags in the different models is based on the Schwarz Bayesian information Criterion (BIC), and all of the lags are kept the same in the subsequent models.

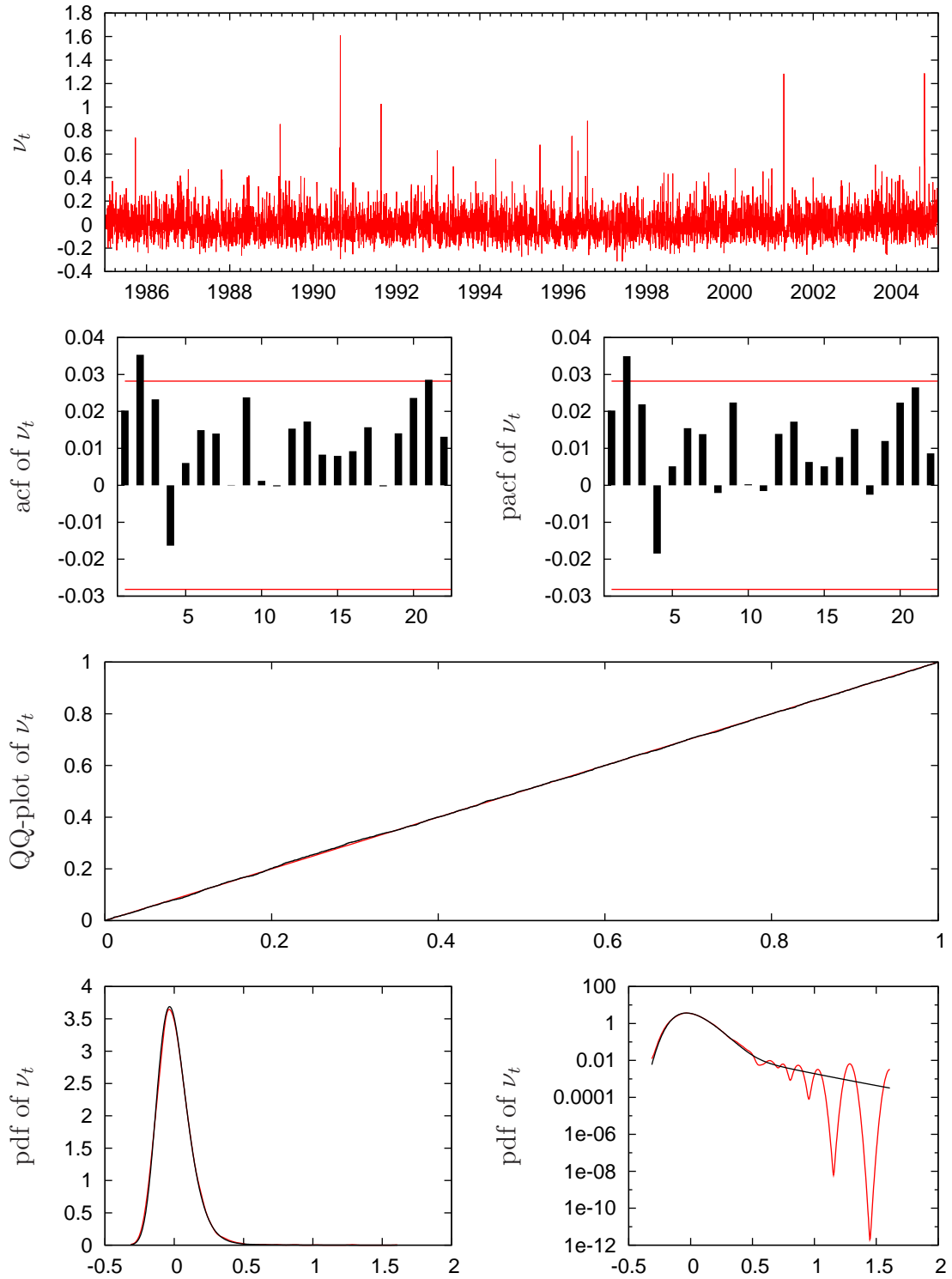
Starting with the results in the first column and the  $BV_t$  equation, the estimates directly mirror earlier results in the literature for the HAR–RV realized volatility model. The daily, weekly and monthly volatility components are all highly statistically significant, while the inclusion of the logarithmic Bipower variation measures over biweekly and other horizons do not improve the fit according to the BIC

**Table 3.2:** Single-Equation Estimation Results of the Auxiliary Model

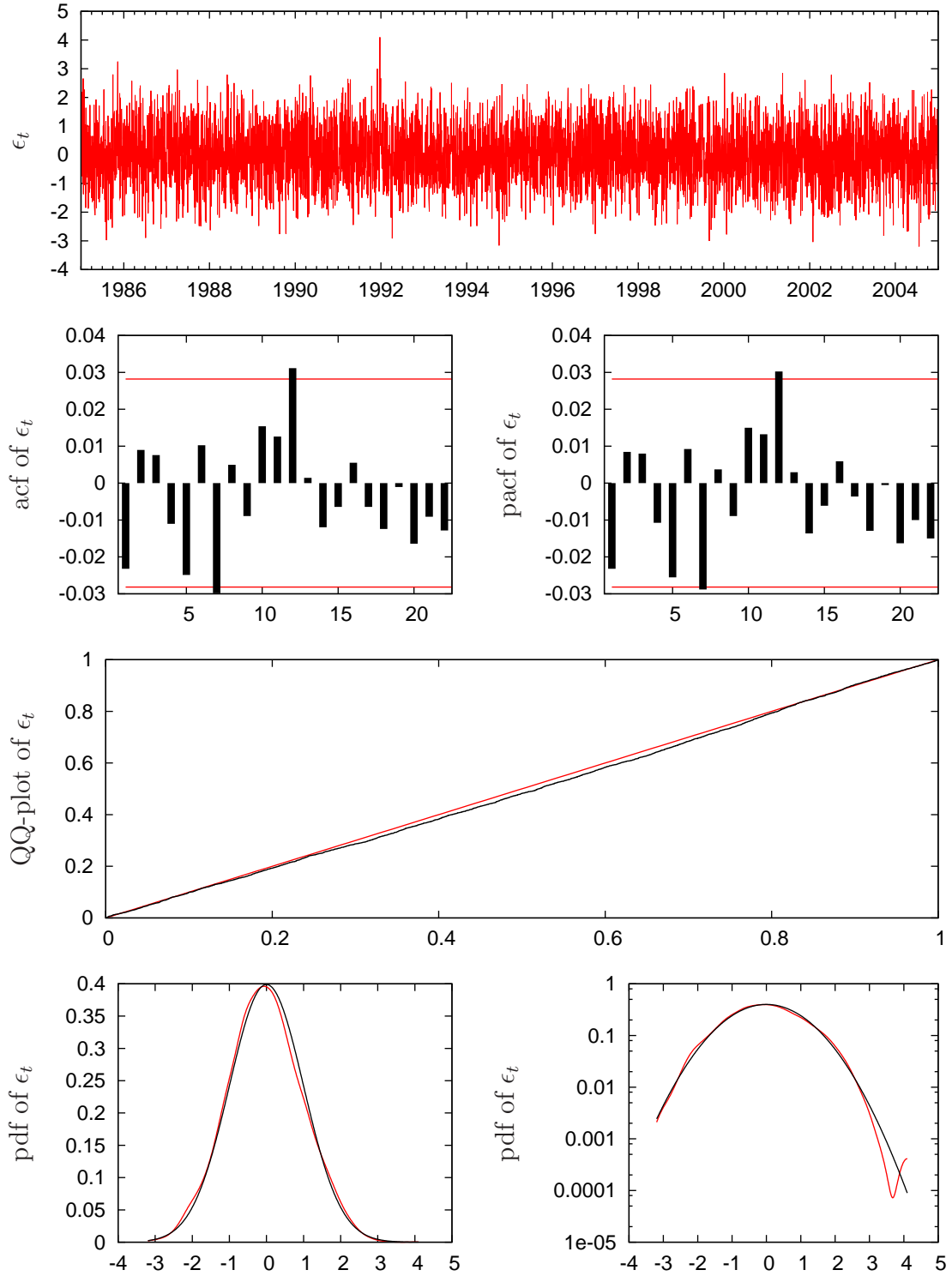
	BV equation			Jump equation			Return equation	
	Estimate	Std. Error		Estimate	Std. Error		Estimate	Std. Error
$\alpha_0$	-0.1978	(0.0170)	$\delta_0$	0.0704	(0.0067)	$\gamma_0$	0.0858	(0.0098)
$\alpha_d$	0.2548	(0.0169)	$\delta_1$	0.0347	(0.0089)	$\gamma_2$	-0.0254	(0.0139)
$\alpha_w$	0.4370	(0.0265)	$\delta_5$	0.0516	(0.0116)	$\gamma_3$	-0.0351	(0.0133)
$\alpha_m$	0.2416	(0.0215)	$\psi_1$	-0.0143	(0.0032)			
$\theta_1$	0.0571	(0.0144)	$\psi_2$	-0.0026	(0.0050)			
$\theta_2$	0.0384	(0.0217)	$\psi_3$	0.0014	(0.0049)			
$\theta_3$	0.1247	(0.0218)	$p_{\nu,2}$	0.0072	(0.0329)			
$\omega$	0.0228	(0.0053)	$\alpha_{NIG}$	71.5659	(52.7253)			
$\alpha_1$	0.0419	(0.0077)	$\beta_{NIG}$	54.0383	(47.7732)			
$\beta_1$	0.8048	(0.0378)	$\delta_{NIG}$	0.2637	(0.0367)			
$p_{u,2}$	0.1451	(0.0304)	$\lambda_{IG}$	0.5247	(0.3198)			
$\mu_{u,2}$	0.7688	(0.1306)	$\mu_{IG}$	1.1804	(5.2968)			
$\sigma_{u,2}$	1.9278	(0.0688)						
logL:	-3464.75		logL:	3775.22		logL:	-5839.63	



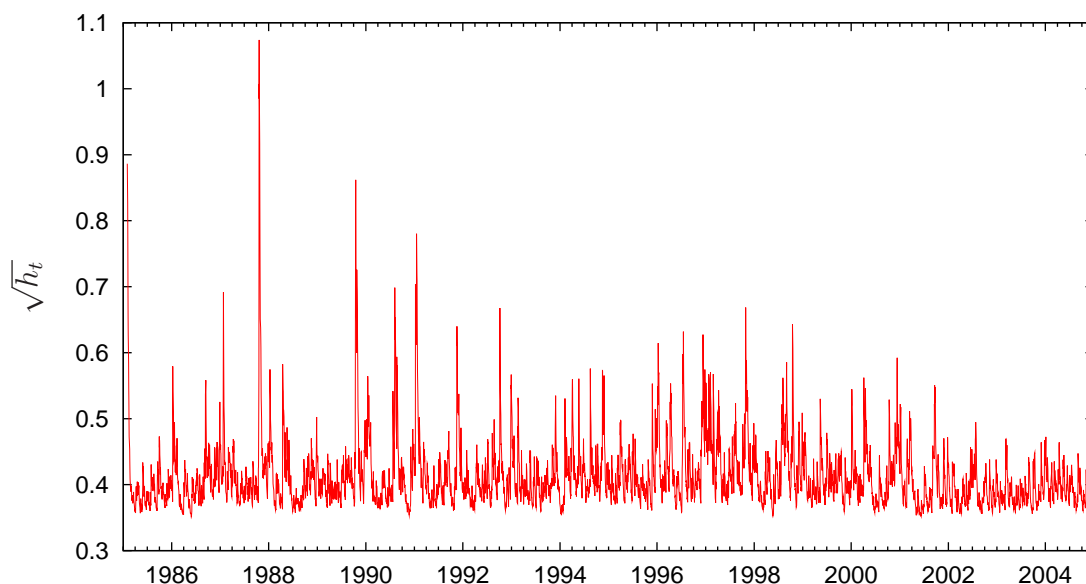
**Figure 3.1:** Residual analysis of the (log.) Bipower variation equation. The upper graph of the figure represents the time evolution of the innovations of the Bipower variation equation. The second line of graphs shows their sample autocorrelations and partial autocorrelations. The third is the corresponding Quantile–Quantile plot. The lower left panel of the figure shows the kernel density estimates of the residuals (red line) and the density of the estimated normal mixture (black line). The right panel shows the same in log scale.



**Figure 3.2:** Residual analysis of the jump equation. The upper graph of the figure represents the time evolution of the innovations of the jump equation. The second line of graphs shows their sample autocorrelations and partial autocorrelations. The third is the corresponding Quantile-Quantile plot. The lower left panel of the figure shows the kernel density estimates of the residuals (red line) and the density of the estimated NIG-IG mixture (black line). The right panel shows the same in log scale.



**Figure 3.3:** Residual analysis of the return equation. The upper graph of the figure represents the time evolution of the innovations of the return equation. The second line of graphs shows their sample autocorrelations and partial autocorrelations. The third is the corresponding Quantile-Quantile plot. The lower left panel of the figure shows the kernel density estimates of the residuals (red line) and the density of a standard normal (black line). The right panel shows the same in log scale.



**Figure 3.4:** The volatility of Bipower variation. The graph exhibits the HAR-GARCH implied volatility series of logarithmic Bipower variation.

criteria. A standard GARCH(1,1) model without any level effects emerges as the preferred specification for the conditional variance. The estimated GARCH parameters easily satisfy the corresponding stationarity condition  $\alpha_1 \sigma_u^2 + \beta_1 < 1$ , where  $\sigma_u^2 = 1 + p_{u,2} (\sigma_{u,2}^2 - 1)$ . The importance of allowing for time-varying volatility is further underscored by the plot of the GARCH conditional standard deviation in Figure 3.4. The importance of asymmetry, or leverage effect, in the continuous volatility component is directly manifest by the highly significant estimates for the  $\theta_1$  and  $\theta_3$  parameters. As expected, the point estimates imply that a lagged negative return shock leads to a much larger increase in the volatility than does a positive shock of the same magnitude. In contrast, the level shift in the volatility equation due to negative news is not significant. This latter result mirrors earlier findings for the realized volatility in Martens et al. (2004). The QQ and kernel density plots in Figure 3.1 indicate that the mixture of two normal distributions does a very good job of capturing the slight skewness and kurtosis inherent in the innovations from the model. Moreover, the autocorrelation and partial autocorrelation functions for the estimated residuals do not reveal any remaining systematic serial correlation within a monthly horizon.

Turning to the jump equation, the autoregressive parameter estimates associated with the first, or daily, and fifth, or weekly, lags are both significant. Still, the magnitude of both coefficients is very small, thus supporting the aforementioned weak own predictability in the jump series. Interestingly, and in sharp contrast to the results for the continuous volatility component, the parameter estimates for  $\psi_2$

and  $\psi_3$  related to the leverage effect suggest that jumps are not asymmetrically affected by lagged return shocks. In fact, if anything the estimate for  $\psi_1$  points to a symmetric, but dampening impact of news on future jumps. The findings of a negative leverage effect in the diffusion volatility component only, is directly in line with most of the parametric jump–diffusion models estimated in the recent literature, in which the leverage effect is typically incorporated by allowing for a negative correlation between the two Brownian motions driving the price and continuous volatility processes; see, e.g., the models in Bates (2000), Eraker et al. (2003) and Pan (2002).<sup>7</sup> Our results on the contemporaneous dependencies in the disturbances, discussed below, further supports this particular specification. The QQ–Plot for the residuals from the  $J_t$  equation as well as the kernel density plots in Figure 3.2 show that the distribution of the jump innovations is well described by the NIG–IG mixture.

The estimates for the return equation in the last column reveal statistically significant, but economically very small, second and third order autocorrelations. As already noted, the standard normal distribution fit the data very well, and are generally preferred over other specifications by the BIC criteria, including a normal with a freely estimated variance parameter as well as a freely estimated zero–mean NIG distribution. We also experimented with the inclusion of a risk premia, or GARCH–in–Mean type effect, by allowing the conditional mean to depend on the realized variance. However, consistent with existing results in the literature suggesting that reliable estimates for this risk premium parameter requires longer return horizons and time–spans of data (e.g. Lundblad, 2004; Ghysels et al., 2005), we found the GARCH–in–Mean effect to be insignificant at the daily level.

### Residual Inter–Dependencies

The separate estimation of the three equations discussed above implicitly assumes that the disturbances are independent. However, based upon existing results in the stochastic volatility literature, we might naturally expect that the disturbances in the return and volatility equations are correlated due to contemporaneous (at the daily level) leverage and/or volatility feedback effects; see, e.g., the recent high–frequency data analysis in Bollerslev et al. (2006b). Moreover, the innovations to the two volatility equations might naturally be expected to be correlated as well. Such inter–dependencies would obviously have to be taken into account in a fully efficient estimation of the joint system, and could also result in inconsistent equation–by–equation estimates.

---

<sup>7</sup>In the context of a representative agent general equilibrium model, Tauchen (2005) has also recently shown that a positive leverage effect can occur depending on the magnitude of the intertemporal marginal rate of substitution and the degree of risk aversion. It is possible that by explicitly differentiating between the two sources of risk, an extension of this model could help explain our empirical findings of a "standard" negative leverage effect in the diffusion component but a positive correlation between returns and jumps.



To begin, consider the sample correlation matrix for the estimated residuals from the Bipower variation, jump and return equations, respectively

$$\hat{\rho} = \begin{bmatrix} 1 & -0.1847 & -0.2008 \\ . & 1 & 0.0283 \\ . & . & 1 \end{bmatrix}.$$

Consistent with the discussion above, the continuous volatility innovations appear to be negatively correlated with both the relative jump residuals and the return innovations. Meanwhile, the correlation between the relative jumps and the return residuals appears negligible.

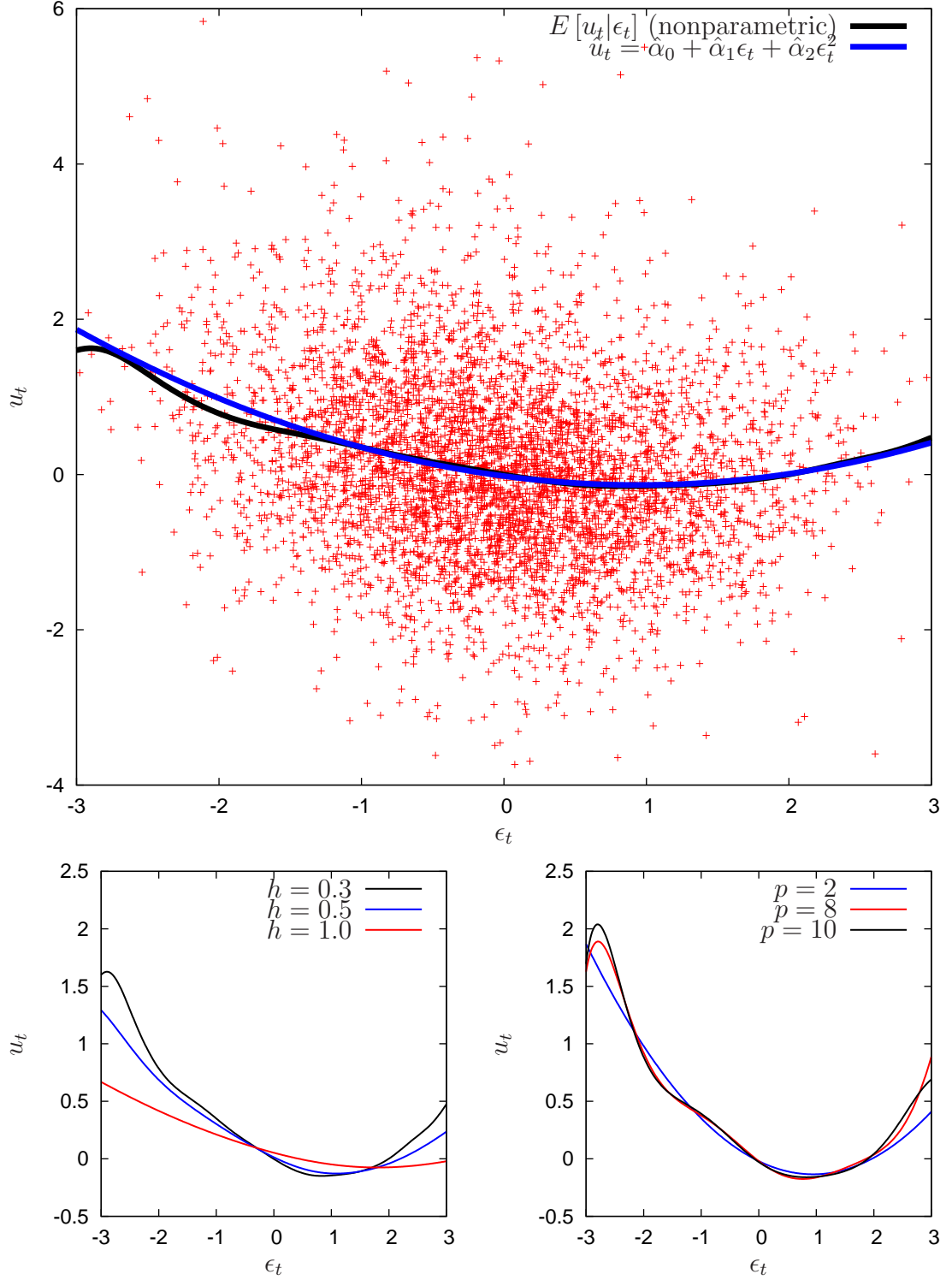
In addition to the linear contemporaneous relationships suggested by the sample correlations, there might also exist nonlinear dependencies due to, e.g., asymmetric volatility effects. To this end, Figures 3.5 to 3.7 present the pairwise scatter plots of the residual series along with a fitted quadratic polynomial, as well as a Rosenblatt–Parzen Gaussian–based kernel estimator. The conjecture of a nonlinear relationship between the residuals is, obviously, supported for at least two of the three combinations. Most obviously, there is an asymmetric negative relation between the residuals of the Bipower equation and the return shocks in Figure 3.5. In fact, this relationship is very similar to the usually assumed lagged leverage effect. In contrast, there is no apparent nonlinear relation between the residuals from the jump and return equations. Interestingly, Figure 3.7 reveals a smirk–like relation between the innovations to the continuous volatility and jump components.<sup>8</sup>

To further visualize the inter–dependencies between the residual series, Figure 3.8 shows the scatter plot of the respective pairwise cumulative distribution functions (cdf). In the absence of any dependencies the points should be uniformly distributed over the whole scatter surface. However, consistent with the aforementioned smile–like pattern in the residual scatter plot for the Bipower variation and return equations, the first panel shows that low (high) cdf values of the return innovations tend to be associated with higher cdf values of the diffusion volatility innovations. A similar pattern emerge in the cdf scatter for the jump and continuous volatility innovations in the bottom panel, but with high return cdf values being associated with smaller values of the jump innovation cdf due to the dampening (smirk–like) behavior. Meanwhile, the cdf scatter between the jump and return innovations in the middle panel exhibits nearly uniformly distributed scatter points.

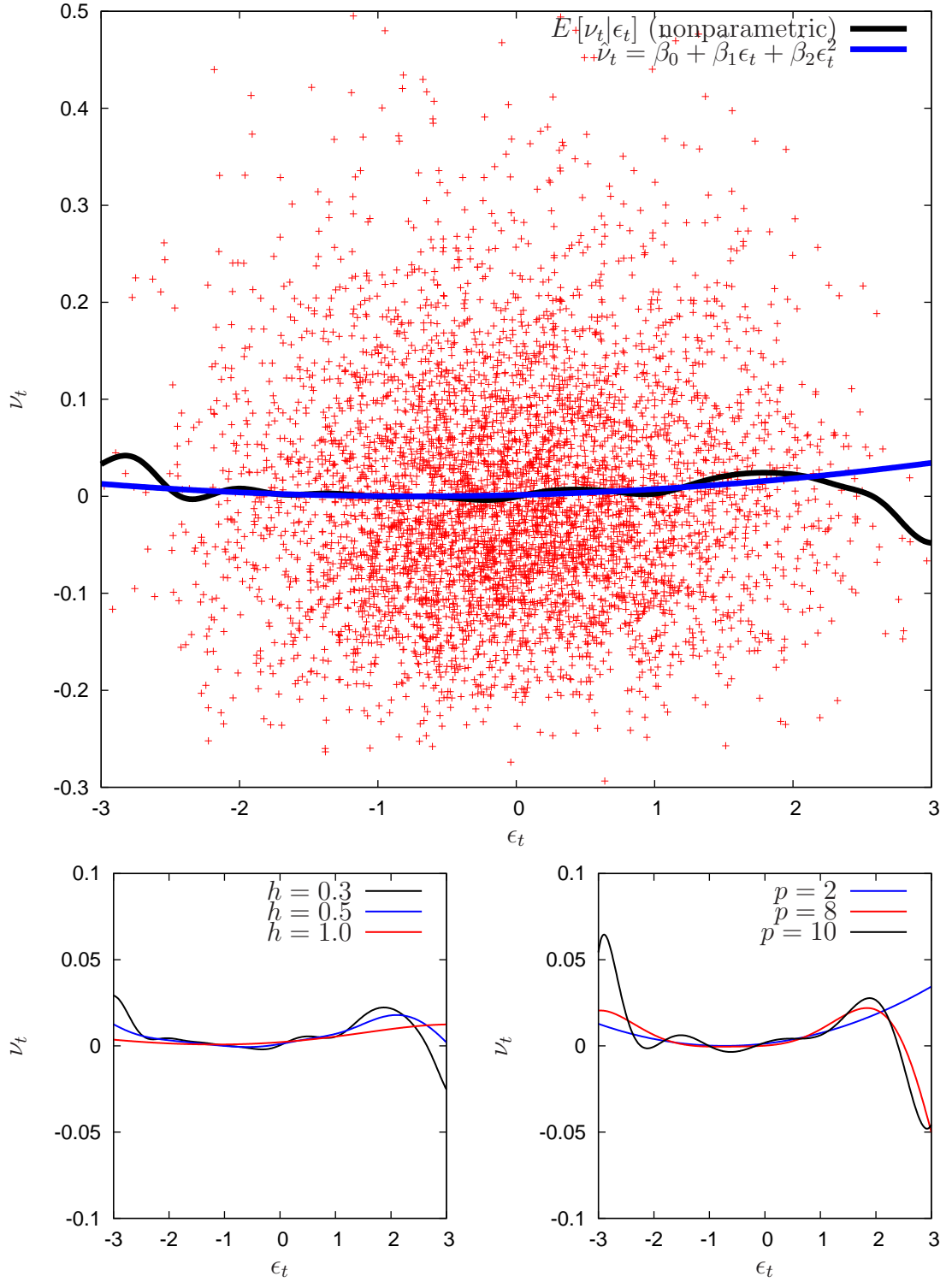
In summary, our analysis points to the existence of important asymmetric dependencies among the three innovation series. These effects should be incorporated into a joint modeling framework in order to, firstly, more systematically quantify and test for their significance, secondly, guard against any biases in the single–equation

---

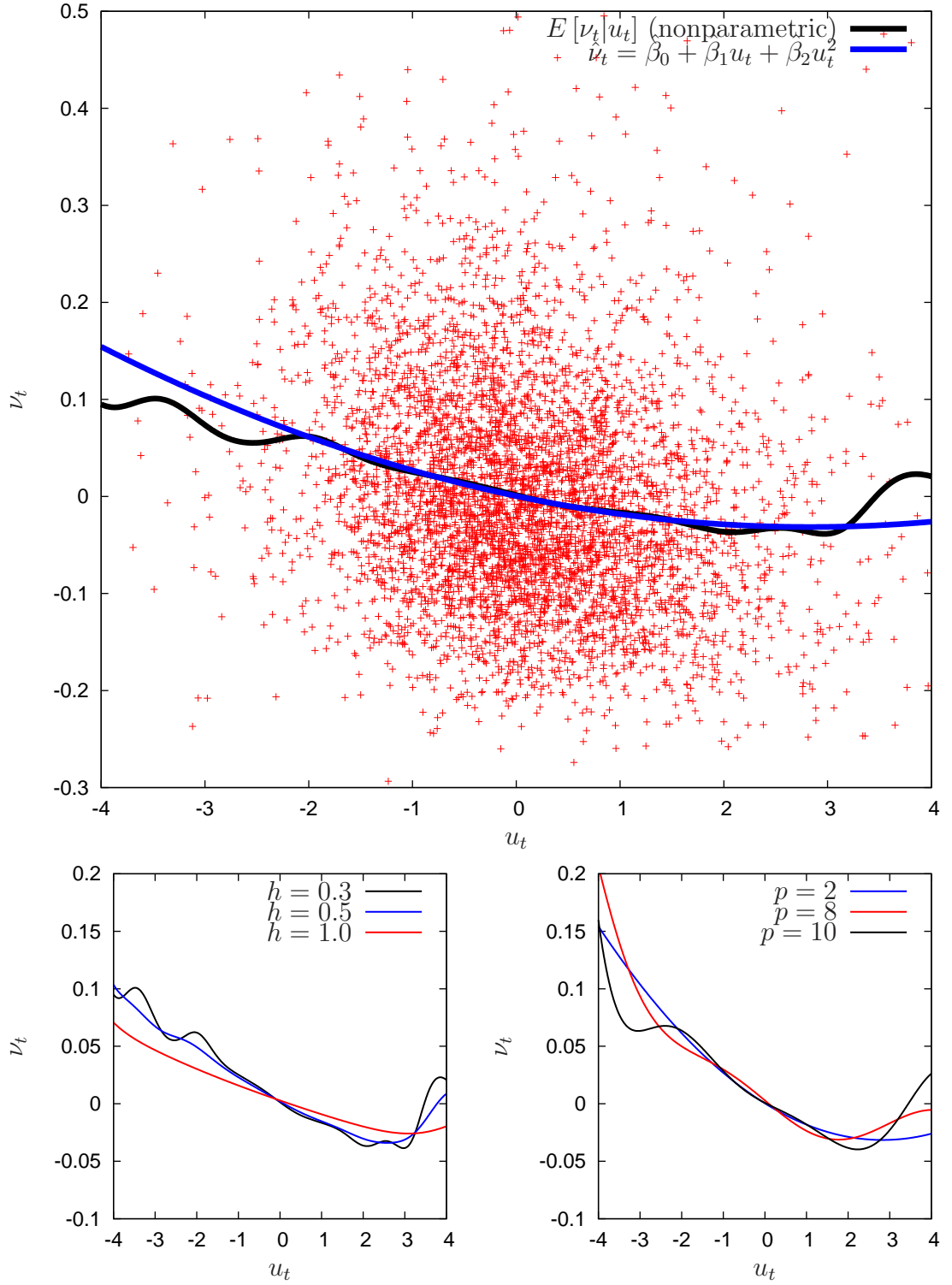
<sup>8</sup>This effect should, of course, be carefully interpreted in light of the definitions of the underlying variation measures. In particular, a negative shock to the (logarithmic) Bipower variation corresponds to an overestimation of the continuous volatility component, which in turn is associated with a larger jump component. In contrast, a positive shock to the Bipower variation equation, and a larger than expected continuous volatility component, does not directly affect the relative jump measure.



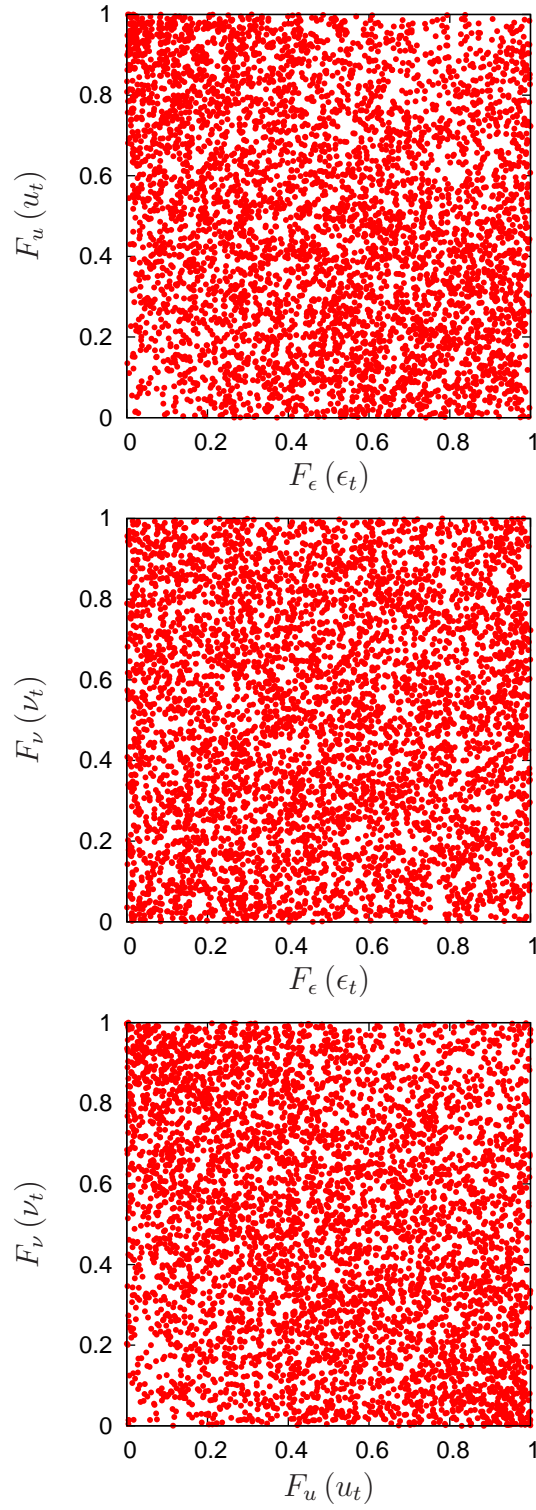
**Figure 3.5:** Dependency analysis of the residuals between the return equation and Bipower variation equation. The lower left and right panels include additional different polynomial and nonparametric specifications, respectively.



**Figure 3.6:** Dependency analysis of the residuals between the return equation and jump equation. The lower left and right panels include additional different polynomial and nonparametric specifications, respectively.



**Figure 3.7:** Dependency analysis of the residuals between the Bipower variation equation and the jump equation. The lower left and right panels include additional different polynomial and nonparametric specifications, respectively.



**Figure 3.8:** CDF scatter plot of the single-equation innovations.

estimates, and thirdly, enhance the efficiency of the individual model parameter estimates. The unified system approach explicitly allowing for nonlinear functional forms of residual dependencies developed in the next section does this.

### 3.3.3 System Estimation

The results of the equation-by-equation estimations suggest that the proposed model specifications provide an adequate description of the dynamic dependencies in the two volatility and return processes, but that it does not fully account for the nonlinear contemporaneous dependencies among the innovations. We therefore retain our basic three equation set up, but additionally model the nonlinear inter-dependencies based on the following system of equations

$$\begin{aligned}
 r_t &= \gamma_0 + \sum_{j=1}^d \gamma_j r_{t-j} + \sqrt{RV_t} \epsilon_t \\
 \log BV_t &= \alpha_0 + \alpha_d \log BV_{t-1} + \alpha_w (\log BV)_{t-5:t-1} + \alpha_m (\log BV)_{t-22:t-1} \\
 &\quad + \theta_1 \frac{|r_{t-1}|}{\sqrt{RV_{t-1}}} + \theta_2 I[r_{t-1} < 0] + \theta_3 \frac{|r_{t-1}|}{\sqrt{RV_{t-1}}} I[r_{t-1} < 0] + \sqrt{h_t} (u_t + g(\epsilon_t)) \\
 h_t &= \omega + \sum_{j=1}^q \alpha_j (\log BV_{t-j} - x'_{BV} \beta_{BV})^2 + \sum_{j=1}^p \beta_j h_{t-j} + \sum_{j=1}^s \lambda_j BV_{t-j} \\
 \log \left( \frac{RV_t}{BV_t} \right) &= \delta_0 + \sum_{j=1}^n \delta_j \log \left( \frac{RV_{t-j}}{BV_{t-j}} \right) \\
 &\quad + \psi_1 \frac{|r_{t-1}|}{\sqrt{RV_{t-1}}} + \psi_2 I[r_{t-1} < 0] + \psi_3 \frac{|r_{t-1}|}{\sqrt{RV_{t-1}}} I[r_{t-1} < 0] \\
 &\quad + (\nu_t + m(u_t) + k(\epsilon_t)).
 \end{aligned} \tag{3.23}$$

In comparison to the individual equations, the system explicitly allows the innovations in the continuous volatility and relative jump equations to depend nonlinearly on the return innovations via the general functions  $g(\epsilon_t)$  and  $k(\epsilon_t)$ , respectively. Similarly, the jump innovations are allowed to depend on the continuous volatility shocks via the  $m(u_t)$  function. Thus, by choosing an adequate functional form for each of these functions, we seek to render the underlying three innovation series to be pairwise independent.

Now, utilizing the recursive structure of the basic model equations along with the contemporaneous independence of the transformed innovations, the transition density for the joint system,  $y_t = (\log BV_t, \log \left( \frac{RV_t}{BV_t} \right), r_t)'$ , is available and can be

readily expressed as

$$\begin{aligned}
 f_y(y_t|x_{t-1};\theta) &= \frac{1}{\sqrt{h_t}\sqrt{RV_t}} \\
 &\times f_\epsilon\left(\underbrace{\frac{r_t - x'_r\beta_r}{\sqrt{RV_t}}}_{\epsilon_t} \middle| \vartheta_\epsilon\right) f_u\left(\underbrace{\frac{\log BV_t - x'_{BV}\beta_{BV}}{\sqrt{h_t}} - g\left(\frac{r_t - x'_r\beta_r}{\exp\{\frac{1}{2}\log RV_t\}}\right)}_{u_t} \middle| \vartheta_u\right) \\
 &\times f_\nu\left(\underbrace{\log\left(\frac{RV_t}{BV_t}\right) - x'_{RV}\beta_{RV} - m(u_t) - k(\epsilon_t)}_{\nu_t} \middle| \vartheta_\nu\right),
 \end{aligned}$$

where as before

$$\begin{aligned}
 \epsilon_t &\stackrel{iid}{\sim} \mathcal{N}(0, 1) \\
 u_t &\stackrel{iid}{\sim} \begin{cases} \mathcal{N}_1(0, 1) & \text{with probability } (1 - p_{u,2}) \\ \mathcal{N}_2(\mu_{u,2}, \sigma_{u,2}^2) & \text{with probability } p_{u,2} \end{cases} \\
 \nu_t &\stackrel{iid}{\sim} \begin{cases} \mathcal{NIG}_0(\alpha_{NIG}, \beta_{NIG}, \delta_{NIG}) & \text{with probability } (1 - p_{\nu,2}) \\ \mathcal{IG}(\lambda_{IG}, \mu_{IG}) & \text{with probability } p_{\nu,2}. \end{cases}
 \end{aligned}$$

To complete the specification, we assume that the nonlinear contemporaneous dependencies among the individual equation innovations may be adequately captured by a set of second order degree polynomials<sup>9</sup>

$$g(\epsilon_t) = g_1\epsilon_t + g_2\epsilon_t^2 \quad (3.24)$$

$$k(\epsilon_t) = k_1\epsilon_t + k_2\epsilon_t^2 \quad (3.25)$$

$$m(u_t) = m_1u_t + m_2u_t^2, \quad (3.26)$$

where for identification purposes we have restricted the three constants to be zero. Fully efficient maximum likelihood estimation of the complete system may now proceed in a standard manner by maximizing the log likelihood function defined by the summation of the logarithmic transition densities over the sample observations.

Comparing the system estimation results reported in Table 3.3 to the equation-by-equation results in Table 3.2, the estimates for most of the individual parameters obviously do not change by much. In particular, our previous conclusions regarding

---

<sup>9</sup>We also experimented with higher order polynomials, but found a simple quadratic sufficient in capturing the smirk-like dependencies over the required range.

**Table 3.3:** System Estimation Results of the Auxiliary Model (logL=-5230.37)

	BV equation			Jump equation			Return equation	
	Estimate	Std. Error		Estimate	Std. Error		Estimate	Std. Error
$\alpha_0$	-0.2526	(0.0172)	$\delta_0$	0.0665	(0.0051)	$\gamma_0$	0.0570	(0.0095)
$\alpha_d$	0.2499	(0.0160)	$\delta_1$	0.0422	(0.0095)	$\gamma_2$	-0.0321	(0.0125)
$\alpha_w$	0.4494	(0.0249)	$\delta_5$	0.0500	(0.0110)	$\gamma_3$	-0.0431	(0.0116)
$\alpha_m$	0.2291	(0.0205)	$\psi_1$	-0.0145	(0.0033)			
$\theta_1$	0.0636	(0.0139)	$\psi_2$	-0.0034	(0.0050)			
$\theta_2$	0.0424	(0.0215)	$\psi_3$	0.0028	(0.0051)			
$\theta_3$	0.1246	(0.0211)	$m_1$	-0.0200	(0.0012)			
$g_1$	-0.2493	(0.0186)	$m_2$	0.0013	(0.0004)			
$g_2$	0.1363	(0.0129)	$k_1$	0.0042	(0.0015)			
$\omega$	0.0250	(0.0055)	$k_2$	0.0018	(0.0011)			
$\alpha_1$	0.0425	(0.0077)	$p_{\nu,2}$	0.0174	(0.0263)			
$\beta_1$	0.7707	(0.0417)	$\alpha_{NIG}$	41.8149	(13.6452)			
$p_{u,2}$	0.1617	(0.0035)	$\beta_{NIG}$	26.1884	(11.2286)			
$\mu_{u,2}$	0.6183	(0.1204)	$\delta_{NIG}$	0.2417	(0.0330)			
$\sigma_{u,2}$	1.9391	(0.0731)	$\lambda_{IG}$	0.3183	(0.0933)			
			$\mu_{IG}$	0.3722	(0.7001)			



the lagged leverage effect in the continuous volatility component and the positive correlation between jump and return innovations all remain intact.<sup>10</sup> Moreover, as expected the asymptotic standard errors for the estimated parameters are generally smaller for the system estimates in Table 3.3, highlighting the gain in (asymptotic) efficiency obtained by jointly estimating the three equations.

Along these lines, the highly significant quadratic term in the  $g(\epsilon_t)$  dependency function clearly indicates that the innovations to the continuous volatility component are nonlinearly related to the innovations to the return equation. In contrast, for the return and jump innovations only  $k_1$  is significant and both of the parameters are numerically very small. The aforementioned nonlinear relationship between the continuous volatility component and the relative jump innovations allowed for by the  $m(u_t)$  dependency function is also strongly supported by the joint estimation. The importance of allowing for contemporaneous nonlinear dependencies among the innovations is further underscored by the likelihood ratio test comparing the fully specified simultaneous equation model to the system equation estimates without the quadratic polynomials, which equals an overwhelmingly significant 597.67.

The model presented in Table 3.3 still includes some individually insignificant parameters. In particular, restricting  $\theta_2 = \psi_2 = \psi_3 = k_2 = 0$ , and re-estimating the model results in a LR test statistic of only 6.619 versus the fully general model. Also, the remaining parameter estimates are hardly affected by restricting these four parameters to be equal to zero. Our final preferred model specification is therefore given by this restricted model in Table 3.4.

As an additional diagnostic check for this final specification, consider the sample correlation between the three residual series

$$\hat{\rho} = \begin{bmatrix} 1 & -0.0221 & -0.0096 \\ . & 1 & -0.0046 \\ . & . & 1 \end{bmatrix}.$$

Compared to the sample correlations for the equation–by–equation residuals reported earlier, these are obviously much closer to zero and generally insignificant. The three scatter plots for the pairwise cdf’s for the system residuals in Figure 3.9 now also appear uniformly distributed over the entire range, indicating that the quadratic polynomials have successfully accounted for the nonlinear contemporaneous dependencies observed in the equation–by–equation residuals.

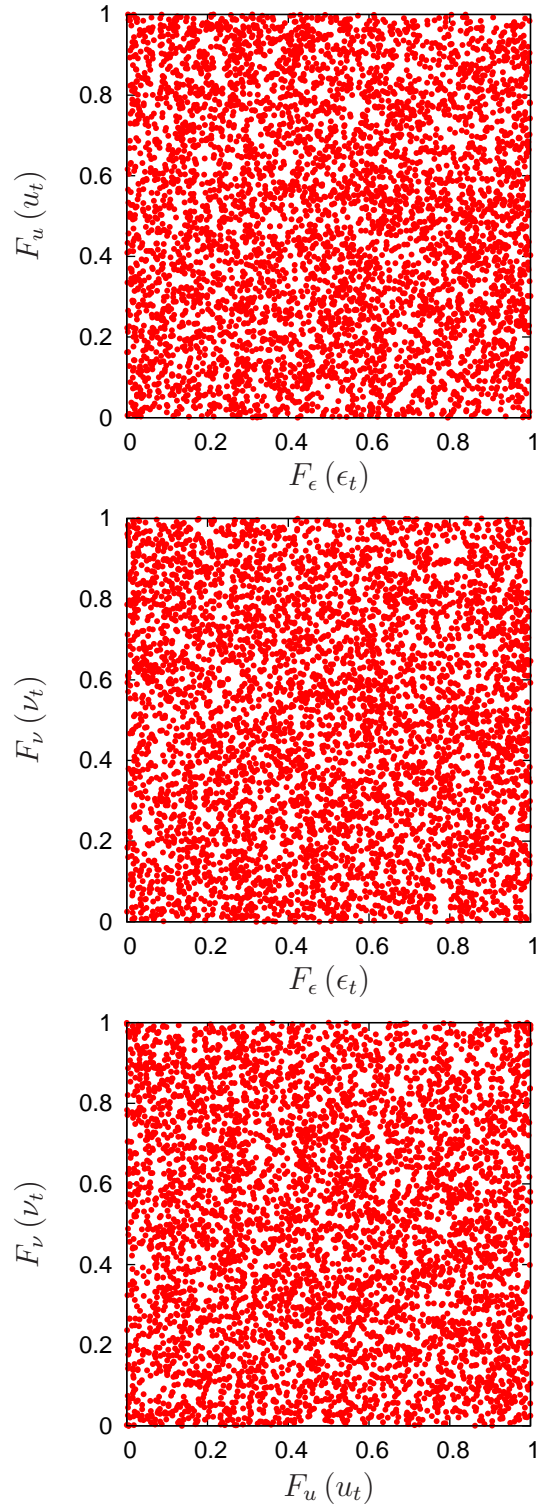
### 3.3.4 Further Accuracy Checks via Simulations

The discussion in the previous section suggests that the model performs an exemplary job in terms of describing the one–day–ahead conditional transition densities when judged by the standard maximum likelihood criteria and corresponding model

<sup>10</sup>Importantly, the system GARCH parameter estimates for the  $BV_t$  equations also satisfy the corresponding second–order stationarity condition:  $\alpha(\sigma_u^2 + g_1^2 + 2g_2^2) + \beta < 1$ , where  $\sigma_u^2 = 1 + p_{u,2}(\sigma_{u,2}^2 - 1)$  and  $\alpha \geq 0, \beta \geq 0$ .

**Table 3.4:** Restricted System Estimation Results of the Auxiliary Model (logL=-5233.67)

	BV equation			Jump equation			Return equation	
	Estimate	Std. Error		Estimate	Std. Error		Estimate	Std. Error
$\alpha_0$	-0.2351	(0.0140)	$\delta_0$	0.0668	(0.0042)	$\gamma_0$	0.0572	(0.0095)
$\alpha_d$	0.2510	(0.0160)	$\delta_1$	0.0426	(0.0094)	$\gamma_2$	-0.0323	(0.0125)
$\alpha_w$	0.4476	(0.0249)	$\delta_5$	0.0497	(0.0110)	$\gamma_3$	-0.0430	(0.0116)
$\alpha_m$	0.2298	(0.0205)	$\psi_1$	-0.0136	(0.0025)			
$\theta_1$	0.0489	(0.0115)	$\psi_2$	-	-			
$\theta_2$	-	-	$\psi_3$	-	-			
$\theta_3$	0.1596	(0.0126)	$m_1$	-0.0200	(0.0012)			
$g_1$	-0.2493	(0.0186)	$m_2$	0.0013	(0.0004)			
$g_2$	0.1406	(0.0127)	$k_1$	0.0045	(0.0015)			
$\omega$	0.0247	(0.0055)	$k_2$	-	-			
$\alpha_1$	0.0419	(0.0077)	$p_{\nu,2}$	0.0198	(0.0236)			
$\beta_1$	0.7728	(0.0416)	$\alpha_{NIG}$	41.0467	(12.6795)			
$p_{u,2}$	0.1628	(0.0355)	$\beta_{NIG}$	25.6054	(10.3213)			
$\mu_{u,2}$	0.6149	(0.1194)	$\delta_{NIG}$	0.2390	(0.0322)			
$\sigma_{u,2}$	1.9374	(0.0730)	$\lambda_{IG}$	0.3007	(0.0830)			
			$\mu_{IG}$	0.3264	(0.5295)			



**Figure 3.9:** CDF scatter plot of the system innovations.

diagnostics.<sup>11</sup> Meanwhile, in order to better understand the workings and possible limitations of a given model, it is often instructive to consider its ability to account for other aspects of the data through the use of simulations. To this end, we generate 105,040 observations from the estimated system, keeping only the last 5,040 observations corresponding to the sample size of our data; i.e., the first 100,000 simulated observations serve as a large burn-in period. We then repeat this 25,000 times, leaving us with 25,000 simulated "daily" sample paths for the returns, logarithmic Bipower variation, and relative jump series. To illustrate, Figure 3.10 shows one such representative set of simulated data. The basic similarities for each of the series with those of the original data in Figure 2.2 are striking, and indeed shows the model to be broadly consistent with the data.

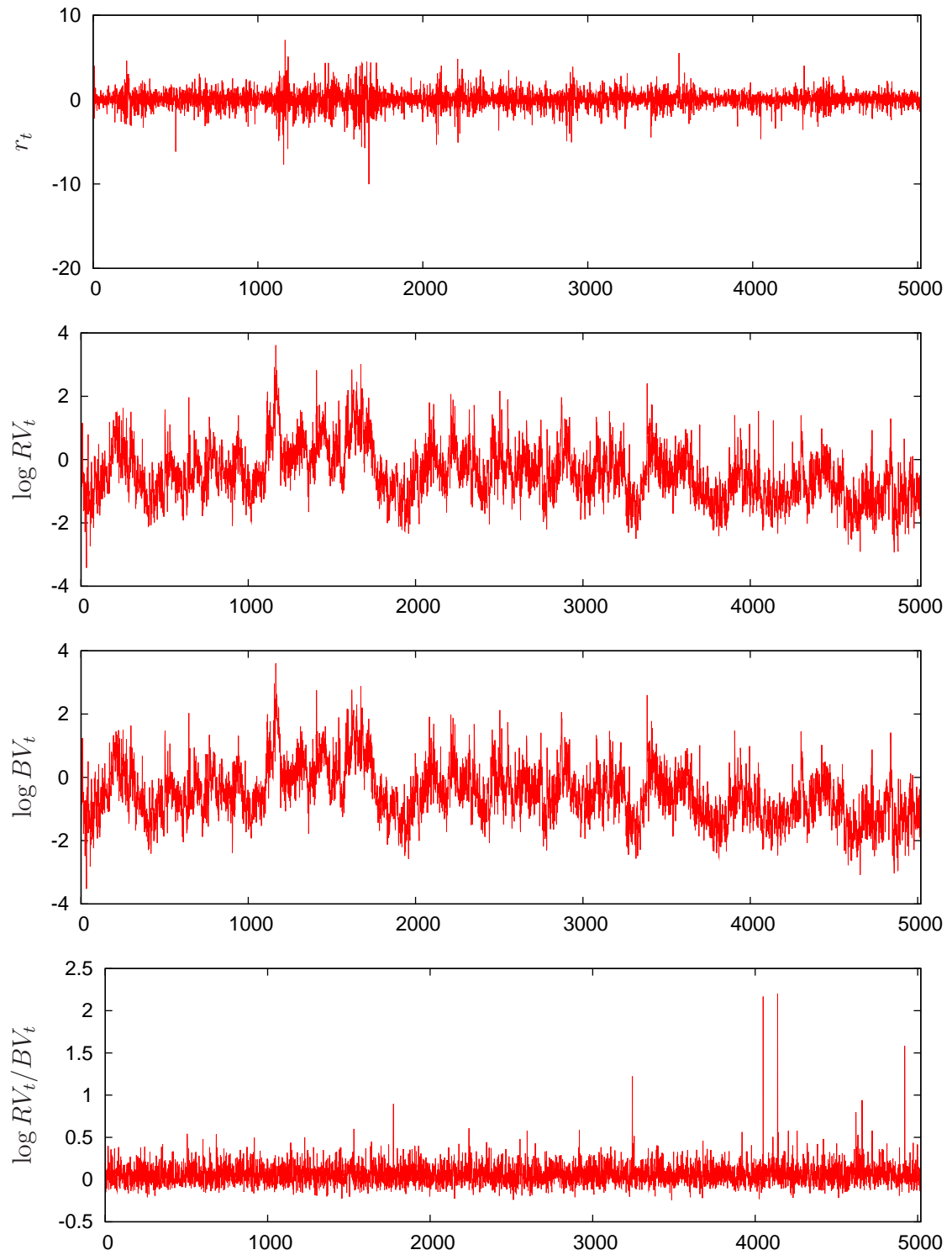
More formally, consider the summary statistics in Table 2.1. By calculating the same set of summary statistics for each of the 25,000 simulated sample paths, we obtain a model-implied sample distribution for the respective statistics. If the model provides an adequate description of the observed data, the realized values of the corresponding sample statistics should lie within reasonable confidence intervals, say 95%, of these model-implied distributions. Table 3.5 provides these 95% simulated confidence intervals for the standard set of summary statistics, as well as the actual sample values from Table 2.1. We also report actual and simulated quantiles for each of the series, and illustrate these in Figure 3.11. Nearly all of the sample statistics, including all of the reported 0.01 to 0.99 quantiles, lie within the simulated confidence bands. Only the realized skewness and kurtosis for the returns and the realized kurtosis for the logarithmic Bipower variation fall outside the 95% bands.<sup>12</sup>

Exploring the dynamic implications of the model, Figure 3.12 shows the sample autocorrelations and partial autocorrelations with the corresponding simulated 95% confidence bands. As can be seen from the figure, the short-run dynamics of both the returns and the relative jump series are generally consistent with those of the model. Meanwhile, the HAR model for  $\log BV_t$ , as well as the model's implications for  $\log RV_t$ , both fall somewhat short in terms of reproducing the highly significant and very slowly decaying sample autocorrelations over longer multi-month lags.<sup>13</sup> At the same time, however, Figure 3.13 shows that the autocorrelations for the Bipower and realized variation expressed in standard deviation form, as would be of interest in many practical applications, both are well accounted for by our

<sup>11</sup> Although not reported in this work, we have also investigated the dynamic patterns and distributional assumptions of the system residuals yielding almost identical results to the ones for the single-equation estimates in Figures 3.1 to 3.3. These results are available upon request.

<sup>12</sup> Although our maximum likelihood based inference doesn't seem to favor this, this could presumably be "fixed" by allowing for a fatter-tailed and skewed error distribution in the return equation, either parametrically or through the use of more flexible semi-nonparametric density estimation as in, e.g., Gallant and Nychka (1987) and Gallant and Tauchen (1989).

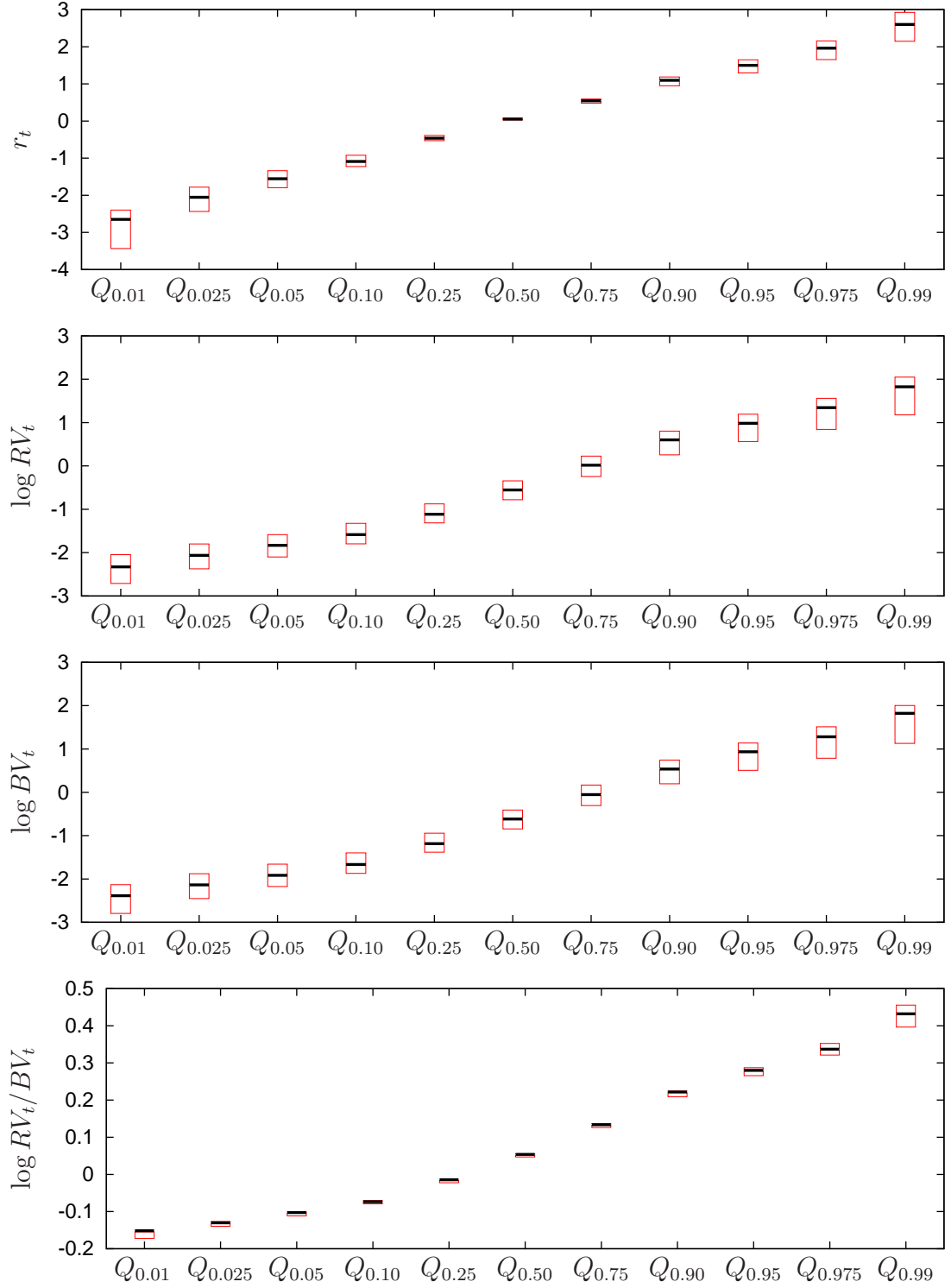
<sup>13</sup> As previously noted, the inclusion of quarterly or longer-run realized variation measures on the right-hand-side of the HAR model for  $\log BV_t$  would presumably remedy this deficiency; see also the simulations reported in Corsi (2004), which shows that the HAR model can get remarkably close to reproducing the autocorrelations of a true long-memory volatility process.



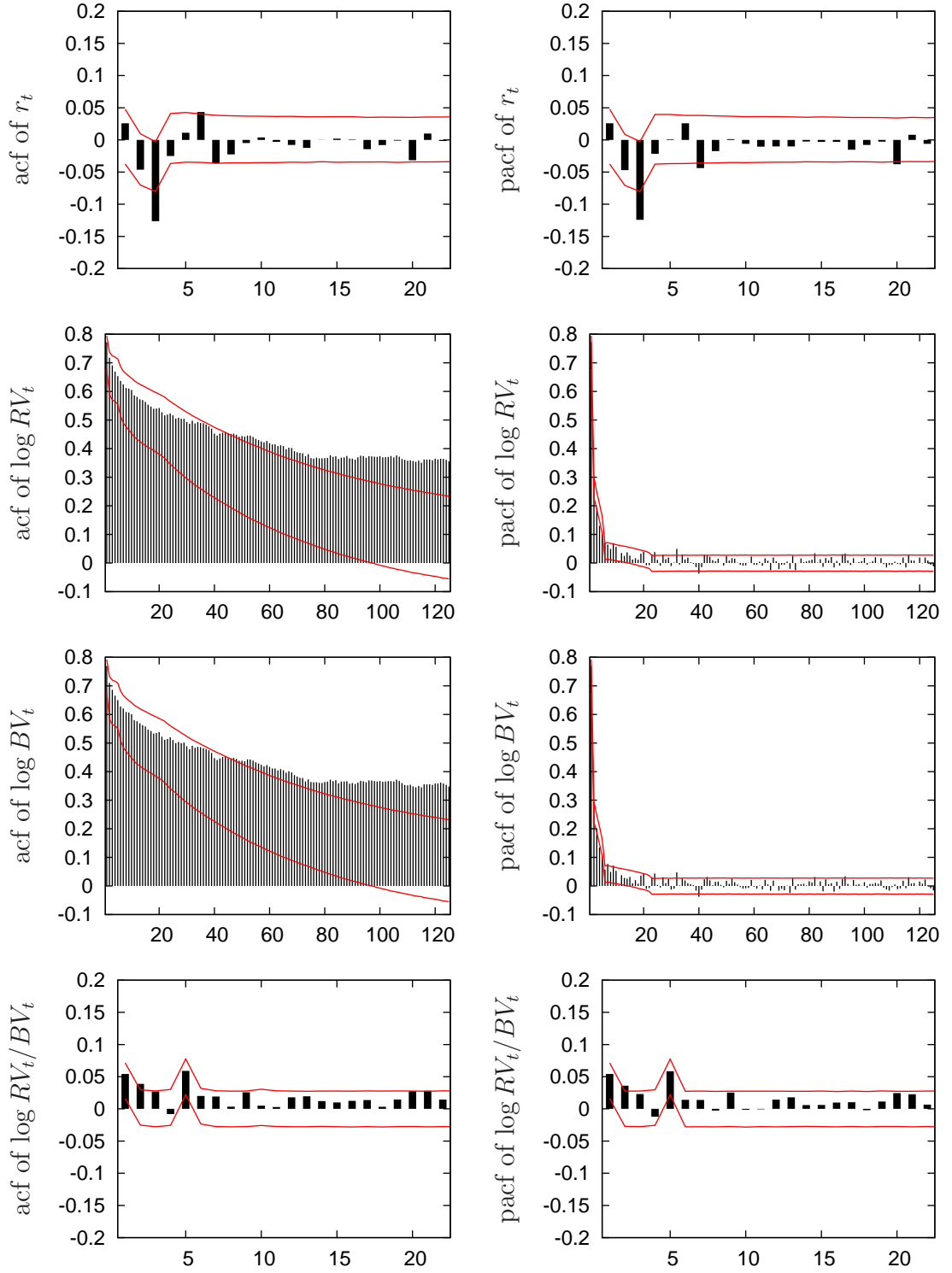
**Figure 3.10:** Simulated paths of returns, logarithmic realized variance, logarithmic Bipower variation and jumps.

Table 3.5: Simulation Results

	$r_t$		$\log RV_t$		$\log BV_t$		$\log \left( \frac{RV_t}{BV_t} \right)$	
stat.	realized	95% intervals	realized	95% intervals	realized	95% intervals	realized	95% intervals
Mean	0.0254	(-0.0125,0.0416)	-0.5139	(-0.7501,-0.3148)	-0.5817	(-0.8159,-0.3788)	0.0678	(0.0611,0.0687)
Std.Dev.	1.0946	(0.8539,1.1504)	0.8775	(0.7382,0.9312)	0.8845	(0.7447,0.9377)	0.1263	(0.1200,0.1354)
Skew.	-2.1648	(-1.8996,0.0906)	0.5948	(-0.0211,0.5765)	0.5416	(-0.0252,0.5711)	1.7761	(0.9838,3.8142)
Exc.Kurt.	96.2483	(3.2207,37.8147)	1.7981	(-0.0476,1.2418)	1.4807	(-0.0354,1.2464)	12.2675	(3.1382,60.7865)
$Q_{0.01}$	-2.6479	(-3.4341,-2.3966)	-2.3275	(-2.7117,-2.0456)	-2.3868	(-2.7956,-2.1305)	-0.1517	(-0.1720,-0.1548)
$Q_{0.025}$	-2.0527	(-2.4330,-1.7798)	-2.0632	(-2.3743,-1.7998)	-2.1377	(-2.4537,-1.8780)	-0.1303	(-0.1394,-0.1268)
$Q_{0.05}$	-1.5535	(-1.7945,-1.3384)	-1.8321	(-2.0994,-1.5834)	-1.9172	(-2.1746,-1.6577)	-0.1027	(-0.1116,-0.1013)
$Q_{0.10}$	-1.0895	(-1.2252,-0.9229)	-1.5848	(-1.7958,-1.3244)	-1.6667	(-1.8696,-1.3969)	-0.0737	(-0.0792,-0.0703)
$Q_{0.25}$	-0.4626	(-0.5275,-0.3872)	-1.1147	(-1.3135,-0.8754)	-1.1859	(-1.3812,-0.9436)	-0.0143	(-0.0225,-0.0147)
$Q_{0.50}$	0.0511	(0.0308,0.0767)	-0.5527	(-0.7797,-0.3452)	-0.6163	(-0.8443,-0.4098)	0.0538	(0.0470,0.0551)
$Q_{0.75}$	0.5446	(0.4790,0.5901)	0.0173	(-0.2421,0.2279)	-0.0533	(-0.3044,0.1683)	0.1339	(0.1264,0.1366)
$Q_{0.90}$	1.0964	(0.9484,1.1850)	0.6022	(0.2553,0.8023)	0.5372	(0.1968,0.7448)	0.2218	(0.2091,0.2242)
$Q_{0.95}$	1.5001	(1.2961,1.6478)	0.9853	(0.5634,1.1937)	0.9352	(0.5084,1.1375)	0.2799	(0.2656,0.2870)
$Q_{0.975}$	1.9627	(1.6544,2.1554)	1.3462	(0.8420,1.5616)	1.2815	(0.7860,1.5113)	0.3371	(0.3211,0.3524)
$Q_{0.99}$	2.5991	(2.1481,2.9251)	1.8250	(1.1784,2.0496)	1.8240	(1.1285,2.0040)	0.4322	(0.3964,0.4557)

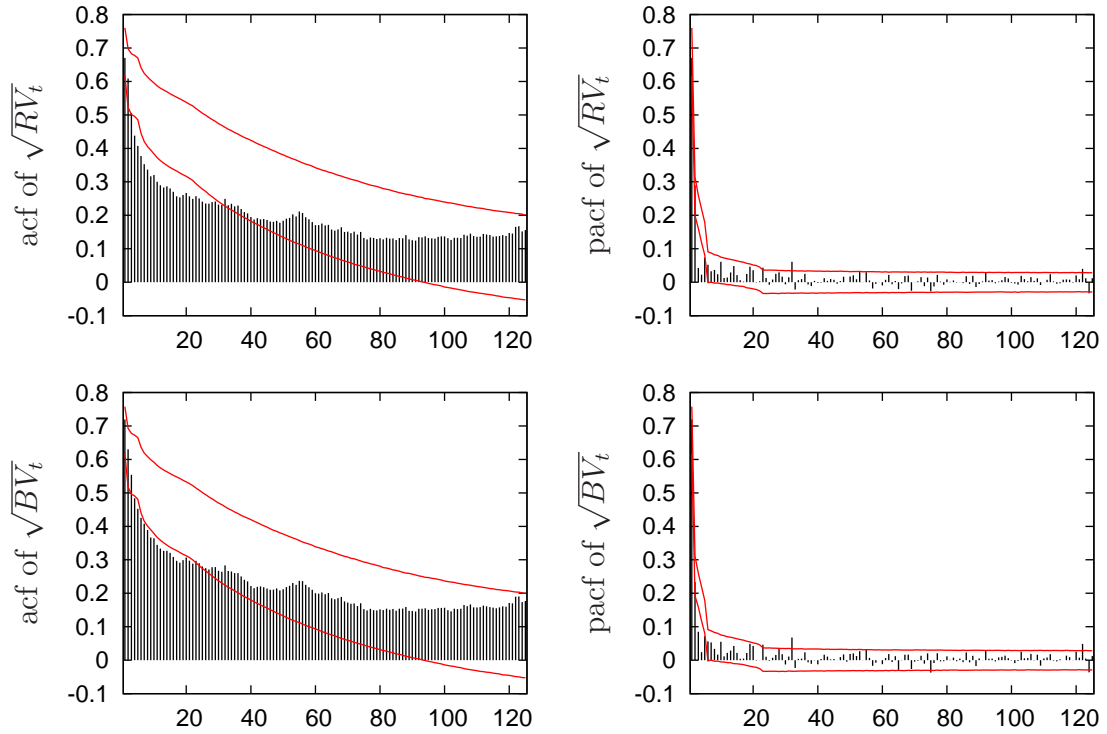


**Figure 3.11:** Sample quantiles of returns, logarithmic realized variance, logarithmic Bipower variation and jumps with 95% simulated confidence intervals.



**Figure 3.12:** Sample autocorrelations and partial autocorrelations of returns, logarithmic realized variance, logarithmic Bipower variation and jumps. The red lines give the upper and lower ranges of the simulated 95% confidence intervals.





**Figure 3.13:** Sample autocorrelations and partial autocorrelations of realized volatility and Bipower variation in standard deviation form. The red lines give the upper and lower ranges of the simulated 95% confidence intervals.

relatively simple and easy-to-implement final preferred model in Table 3.4.

Our estimation and simulation results show, that our simultaneous equation model provides indeed a highly accurate description of the discrete-time joint dynamics of the returns and the two volatility components. Moreover, based on the availability of its transition density, the model can be used in the indirect estimation—e.g. via GSM—of other parametric volatility models, such as the affine or logarithmic stochastic volatility models, while effectively incorporating the relevant information contained in the high-frequency data.

### 3.4 Prior Information

Recall that in contrast to other indirect inference approaches, the GSM method also allows to incorporate some prior information on both, parameters and functionals of the structural model as well as of the auxiliary model.

In our auxiliary model we impose the following prior information. To assure that the auxiliary model always satisfies the general stationarity conditions we impose some parameter restrictions in form of support conditions. In particular

in the Bipower variation equation we restrict the autoregressive component by  $|\alpha_d + \alpha_w + \alpha_m| < 1$ , and require the GARCH parameters to fulfill  $\alpha(\sigma_u^2 + g_1^2 + 2g_2^2) + \beta < 1$ , where  $\sigma_u^2 = 1 + p_{u,2}(\sigma_{u,2}^2 - 1)$  and  $\alpha \geq 0$ ,  $\beta \geq 0$ , assuming a GARCH(1,1) specification. Similar conditions are imposed on the autoregressive parameters of the jump equation, i.e.  $|\sum_{j=1}^n \delta_j| < 1$ . Obviously, the persistence in the return equation also has to be less than one, but we make here additional use of the results reported in the literature stating that returns are only weakly autocorrelated, i.e. they are less predictable. We incorporate this information in form of the following normal priors  $P(|\gamma_j - 0| < 0.098) < 0.95$  for  $j = 1, \dots, d$ , which is part of  $\pi(\eta, v)$ .

Moreover, since mixtures of distributions are notoriously difficult to estimate, we will fix the parameters of the underlying distributions (i.e.  $\mu_{u,2}$ ,  $\sigma_{u,2}$  and  $\alpha_{NIG}$ ,  $\beta_{NIG}$ ,  $\delta_{NIG}$ ,  $\delta_{IG}$ ,  $\mu_{IG}$ ) to the estimated values from our observed data (see Table 3.4), and only allow the mixing probabilities to be freely estimated within the GSM routine. In addition, recall from (3.16) and (3.19) that the mixing probabilities in each mixture of distributions, are restricted to sum to one.

For the structural models we only impose the normalizations discussed in Section 3.1 to achieve identification as well as non–explosivity restrictions. Furthermore, we do not consider other prior information, in order to allow for comparability among the different competing continuous–time stochastic volatility models, i.e. the assessment of each of the models should be subject to the same information set. Note, that in contrast the prior information on the auxiliary model is not relevant only for one specific structural model, but for all of the continuous–time models which are estimated with respect to this prior information.

### 3.5 Empirical Results

Given the prior information and the auxiliary model, we can now turn to the estimation of the continuous–time stochastic volatility models using GSM. In the following we provide a discussion of the corresponding estimation results. Importantly, note that to this end we conduct our model assessment only in terms of the auxiliary model, although, of course, the parameter estimates of the structural models are also available as in GSM there exists for each parameter vector of the structural model an associated parameter vector of the auxiliary model with the correspondence given by the map. However, analyzing the results in terms of the auxiliary model facilitates the comparison of the different continuous–time stochastic volatility models among each other, as well as the assessment of their ability to reproduce the stylized facts of the data, which is the main objective of this study.

Table 3.6 presents the corresponding estimation results of the continuous–time stochastic volatility models along with the parameter estimates of the auxiliary model which are just reproduced from Table 3.2. Note, that rather than testing for the significance of the individual parameters of the continuous–time stochastic volatility models, we are mainly interested in the closeness of these estimates to the ones of the auxiliary model. We therefore only report the standard deviations of the

**Table 3.6:** Estimation Results of the Continuous-Time Stochastic Volatility Models

	Param.	Est.	Stdev.	LL1V	AFF1V	AFF1V-J-75	AFF2V
<i>BV</i>	$\alpha_0$	-0.235	0.014	-0.107	-0.101	0.026	0.011
	$\alpha_d$	0.251	0.016	0.833	0.816	0.210	0.325
	$\alpha_w$	0.448	0.025	-0.067	-0.016	0.694	0.590
	$\alpha_m$	0.230	0.021	-0.004	-0.003	0.079	0.074
	$\theta_1$	0.049	0.012	-0.056	-0.056	-0.060	-0.044
	$\theta_3$	0.160	0.013	0.119	0.122	0.113	0.095
	$p_{u,2}$	0.163	0.036	0.001	0.021	0.012	0.012
	$\omega$	0.025	0.006	0.032	0.060	0.001	0.001
	$\beta_1$	0.773	0.042	0.889	0.689	0.929	0.940
	$\alpha_1$	0.042	0.008	0.001	0.072	0.051	0.047
<i>r</i>	$\gamma_0$	0.057	0.010	0.055	0.058	0.044	0.078
	$\gamma_2$	-0.032	0.013	-0.001	-0.007	-0.021	-0.007
	$\gamma_3$	-0.043	0.012	0.002	0.002	-0.014	-0.018
<i>J</i>	$\delta_0$	0.067	0.004	0.026	0.029	0.034	0.028
	$\delta_1$	0.043	0.009	0.013	0.024	-0.006	0.061
	$\delta_5$	0.050	0.011	-0.018	-0.006	-0.008	0.014
	$\psi_1$	-0.014	0.003	0.001	0.001	-0.002	-0.000
	$p_{\nu,2}$	0.020	0.024	0.000	0.000	0.082	0.000
contemp.	$g_1$	-0.249	0.019	-0.146	-0.152	-0.128	-0.106
	$g_2$	0.141	0.013	0.007	-0.003	-0.004	0.007
	$m_1$	-0.020	0.001	-0.014	-0.016	-0.023	-0.029
	$m_2$	0.001	0.000	0.000	-0.003	0.001	-0.000
	$k_1$	0.005	0.002	-0.000	-0.003	0.001	-0.001

parameter estimates of the auxiliary models providing us with confidence intervals, which should include the estimates of the continuous–time models if those models are adequate.

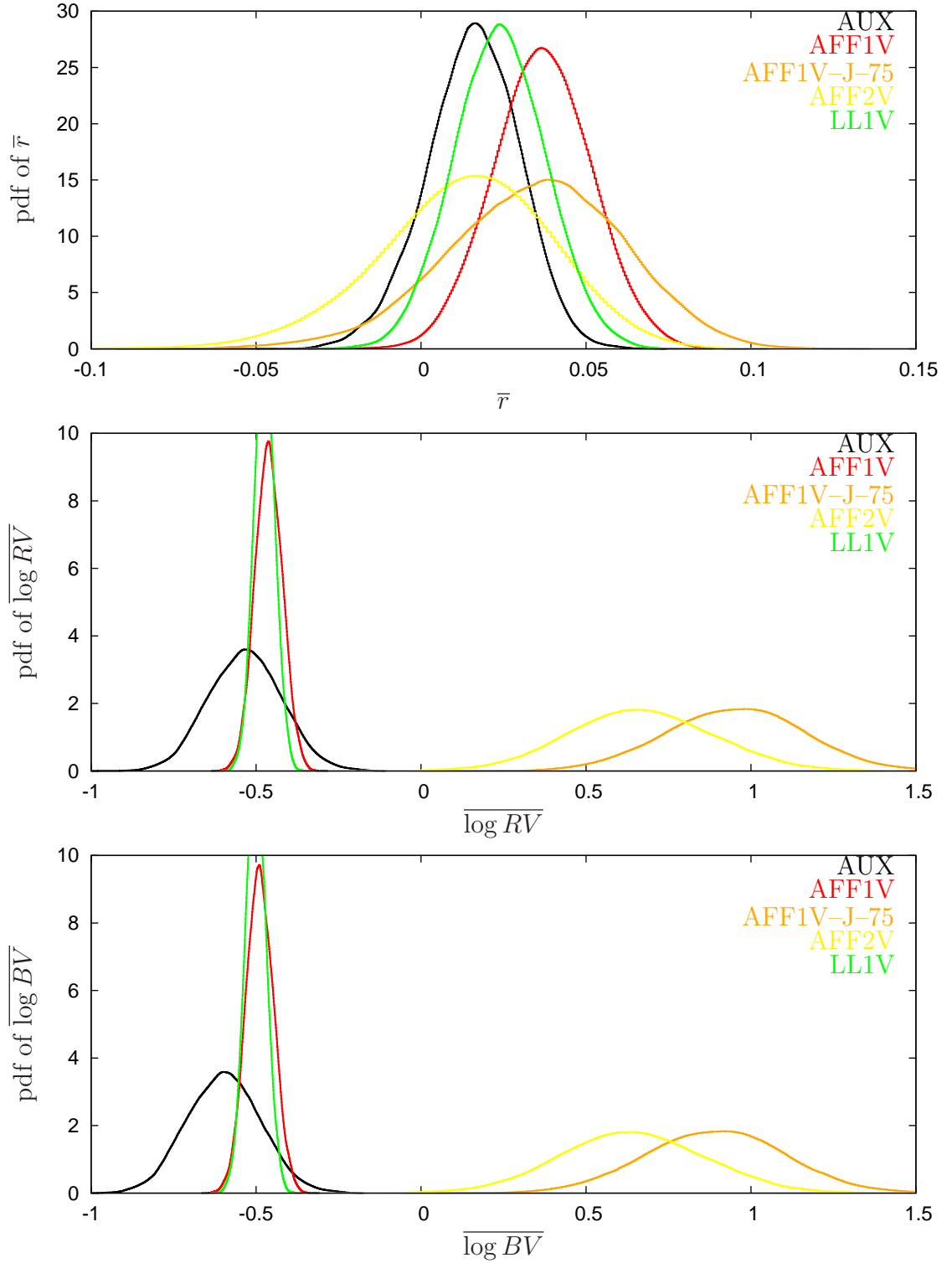
Let us first focus on the parameters of the Bipower variation equation (first panel). Comparing the estimates of the HAR coefficients reveals that the one–factor models introduce persistence mainly through the daily volatility component. In particular, they strongly overstate the impact of the daily volatility component ( $\alpha_d$ ) and impose nearly zero weight to the weekly and monthly components. In contrast, the two–factor model and the affine one–factor model with jumps attribute the highest impact to the weekly component followed by the daily one, which is just in line with the data, i.e. the auxiliary model, where, however, the weekly and monthly coefficients are smaller and larger, respectively. Furthermore, overall the AFF1V and the AFF1V-J-75 overestimate the persistence as measured by the sum of the HAR coefficients, whereas the one–factor models underestimate it.<sup>14</sup> Turning our attention to the parameters of the leverage specification, i.e.  $\theta_1$  and  $\theta_3$ , it becomes obvious that none of the continuous–time models is able to reproduce the smile–like relationship between returns and volatility. The estimates rather imply a smirk–like behavior as positive lagged returns lead to a decrease in the continuous–volatility component rather than an increase, while negative returns still have a positive impact. Similarly, the contemporaneous interdependencies between the Bipower variation and the return innovations is not captured. The results for the return equation show, that the weak serial correlation in the returns is not generated by any of the considered continuous–time stochastic volatility models as is also the case for the own serial correlation of the jumps, which, however, is not surprising, given that the model specifications impose no serial dependence in the jump process. Moreover, returns seem to have no lagged nor contemporaneous impact on the jumps as is indicated by the near zero estimates of  $\psi_1$  and  $k_1$ , respectively. However, in the affine one–factor model with jumps the innovations exhibit similar dependency on the contemporaneous shocks in the Bipower variation as is observed empirically via the auxiliary model. Importantly, the model with jumps is also the only one that has a positive mixing probability  $p_{\nu,2}$ . However, according to the estimated value, the model implies a larger relative importance of the jumps than is observed empirically.

Overall, none of the continuous–time stochastic volatility models seems to be able to reproduce all of the empirical characteristics of our data. However, some of the data characteristics cannot be attributed to particular parameters of the auxiliary model, and we therefore supplement our analysis by a simulation–based unconditional and conditional model assessment.

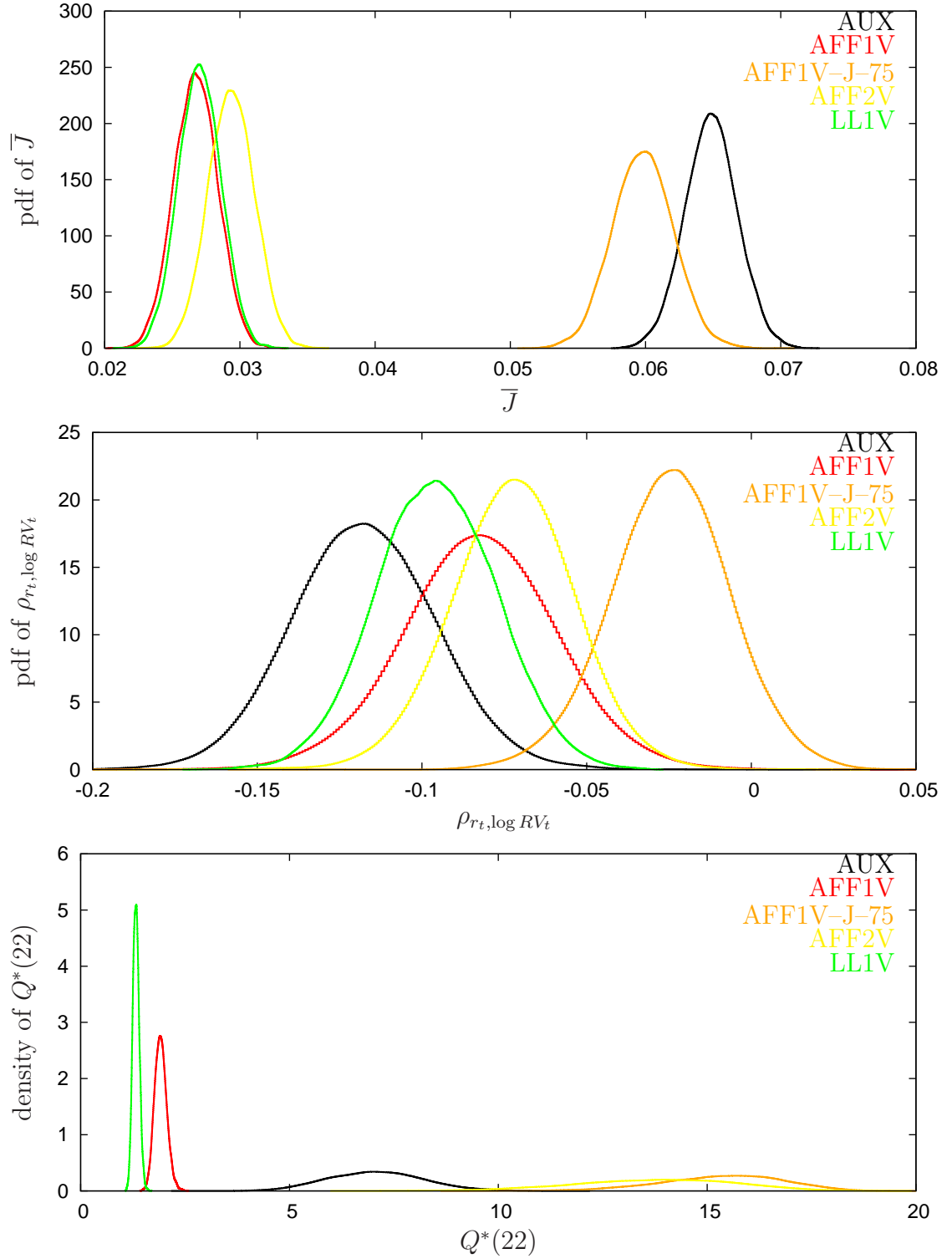
---

<sup>14</sup>Note, that AFF1V-J-75 denotes the affine one–factor stochastic volatility model with  $\lambda = 75$ .

In particular, we have fixed  $\lambda$  to different values and have reestimated the AFF1VJ model for each of these values whereby it turned out that on average 75 jumps per year are most adequate. We henceforth fixed  $\lambda = 75$ .



**Figure 3.14:** The figure shows the model-implied distributions of the means of the returns (upper panel), of the logarithmic realized variance (middle panel), and of the logarithm of Bipower variation (lower panel). The distributions are based on a simulation study, which is further described in Section 3.5. For an explanation of the model abbreviations see Table 3.1.



**Figure 3.15:** The figure shows the model–implied distributions of the mean of the jump measure (upper panel), of the correlation between returns and logarithmic realized variance (middle panel), and of the Ljung–Box statistics (scaled by the number of observations) on serial correlation of realized variance up to order 22 (lower panel). The distributions are based on a simulation study, which is further described in Section 3.5. For an explanation of the model abbreviations see Table 3.1.

**Table 3.7:** Summary Statistics of Model-Implied Distributions

model	mean	std	$Q_{0.05}$	$Q_{0.25}$	$Q_{0.50}$	$Q_{0.75}$	$Q_{0.95}$
$\bar{r}$							
AUX	0.0157	0.0140	-0.0083	0.0067	0.0153	0.0255	0.0378
AFF1V	0.0363	0.0146	0.0124	0.0267	0.0355	0.0462	0.0603
AFF1V-J-75	0.0338	0.0270	-0.0115	0.0163	0.0334	0.0524	0.0761
AFF2V	0.0132	0.0256	-0.0312	-0.0021	0.0132	0.0307	0.0519
LL1V	0.0236	0.0138	0.0010	0.0142	0.0227	0.0329	0.0463
$\overline{\log RV}$							
AUX	-0.5325	0.1101	-0.7108	-0.6087	-0.5392	-0.4579	-0.3514
AFF1V	-0.4626	0.0411	-0.5299	-0.4906	-0.4656	-0.4345	-0.3947
AFF1V-J-75	0.9507	0.2107	0.6036	0.8078	0.9392	1.0934	1.2951
AFF2V	0.6666	0.2198	0.3107	0.5179	0.6500	0.8138	1.0308
LL1V	-0.4750	0.0316	-0.5263	-0.4964	-0.4774	-0.4533	-0.4226
$\overline{\log BV}$							
AUX	-0.5973	0.1105	-0.7761	-0.6736	-0.6042	-0.5226	-0.4154
AFF1V	-0.4893	0.0413	-0.5571	-0.5175	-0.4924	-0.4612	-0.4212
AFF1V-J-75	0.8909	0.2111	0.5433	0.7470	0.8792	1.0341	1.2350
AFF2V	0.6372	0.2202	0.2794	0.4882	0.6204	0.7846	1.0014
LL1V	-0.5020	0.0318	-0.5536	-0.5236	-0.5045	-0.4802	-0.4495
$\bar{J}$							
AUX	0.0648	0.0019	0.0617	0.0636	0.0647	0.0661	0.0681
AFF1V	0.0268	0.0016	0.0241	0.0256	0.0266	0.0278	0.0294
AFF1V-J-75	0.0598	0.0022	0.0562	0.0583	0.0597	0.0613	0.0635
AFF2V	0.0293	0.0017	0.0266	0.0282	0.0292	0.0305	0.0321
LL1V	0.0270	0.0016	0.0245	0.0260	0.0269	0.0281	0.0296
$\rho_{r_t, \log RV_t}$							
AUX	-0.1183	0.0215	-0.1528	-0.1325	-0.1197	-0.1042	-0.0835
AFF1V	-0.0833	0.0229	-0.1206	-0.0981	-0.0846	-0.0686	-0.0469
AFF1V-J-75	-0.0252	0.0181	-0.0553	-0.0369	-0.0260	-0.0132	0.0043
AFF2V	-0.0724	0.0187	-0.1028	-0.0843	-0.0733	-0.0604	-0.0427
LL1V	-0.0956	0.0182	-0.1254	-0.1080	-0.0969	-0.0832	-0.0656
$Q^*(22)$							
AUX	7.0609	1.1347	5.2424	6.2495	6.9712	7.8218	8.9628
AFF1V	1.9170	0.1444	1.6945	1.8164	1.9024	2.0098	2.1641
AFF1V-J-75	15.3082	1.4748	12.7166	14.3608	15.3243	16.3591	17.5469
AFF2V	13.6959	1.9287	10.3657	12.3940	13.6654	15.0849	16.6846
LL1V	1.3268	0.0765	1.2061	1.2733	1.3195	1.3764	1.4551

*Continued on next page*

model	mean	std	$Q_{0.05}$	$Q_{0.25}$	$Q_{0.50}$	$Q_{0.75}$	$Q_{0.95}$
$\sigma_r$							
AUX	0.9901	0.0792	0.8782	0.9370	0.9795	1.0349	1.1211
AFF1V	1.0167	0.1113	0.9565	0.9851	1.0065	1.0358	1.0863
AFF1V-J-75	1.9088	0.2329	1.5698	1.7470	1.8771	2.0471	2.3072
AFF2V	1.6334	0.2642	1.3239	1.4762	1.5873	1.7500	2.0433
LL1V	0.9691	0.0233	0.9324	0.9529	0.9667	0.9848	1.0086
$\sigma_{\log RV}$							
AUX	0.8332	0.0487	0.7584	0.7988	0.8276	0.8641	0.9176
AFF1V	0.9552	0.0254	0.9138	0.9380	0.9530	0.9718	0.9972
AFF1V-J-75	0.7877	0.1058	0.6308	0.7125	0.7727	0.8545	0.9752
AFF2V	0.7314	0.1145	0.5724	0.6507	0.7104	0.7959	0.9362
LL1V	0.8955	0.0172	0.8676	0.8838	0.8941	0.9071	0.9243
$\sigma_{\log BV}$							
AUX	0.8267	0.0487	0.7520	0.7926	0.8212	0.8576	0.9112
AFF1V	0.9521	0.0252	0.9110	0.9352	0.9500	0.9686	0.9938
AFF1V-J-75	0.7974	0.1042	0.6429	0.7235	0.7822	0.8630	0.9823
AFF2V	0.7294	0.1142	0.5702	0.6492	0.7084	0.7939	0.9331
LL1V	0.8936	0.0171	0.8659	0.8820	0.8922	0.9050	0.9222

Notes: The table reports the summary statistics, i.e. the mean, standard deviation and the different quantiles, of the model–implied distributions of different statistics. Note, that **bar** denotes the mean of the respective series,  $\rho$  the correlation coefficient,  $Q^*(22)$  is the Ljung–Box statistics (scaled by the number of observations) for serial correlation up to order 22, and  $\sigma$  denotes the standard deviation of the series indicated in the subscript. The distributions are based on a simulation study, which is further described in Section 3.5. For an explanation of the model abbreviations see Table 3.1.

Our unconditional model assessment is conducted as follows. We simulate from our estimated continuous–time stochastic volatility models as well as from the fitted auxiliary model a series of daily returns from which we construct a series of 15,040 daily returns, Bipower variation, realized variance, and of the logarithmic relative jump measure. The first 10,000 observations serve as a large burn–in period. From the remaining 5,040 observations—which is equivalent to the size of our sample—we compute different statistics, e.g. the mean of the jump measure. Repeating this procedure 10,000 times provides us with a model–implied distribution of the respective statistics, which are partly presented in Figures 3.14 and 3.15. Table 3.7 reports some summary statistics of all the distributions considered. Figure 3.14 and the corresponding first three panels of Table 3.7 show that the means of the logarithmic realized variance and logarithmic Bipower variation are best captured by the one–factor stochastic volatility models, whereby, however, their implied distribution exhibits much less variance than the empirical one as given by the distribution from the simulation based on the auxiliary model. In contrast, the two–factor model and the one–factor model with jumps imply considerably higher means of the two series. This in turn leads to a higher variance in the return distribution, whereby its mean is more or less reproduced by all of the different stochastic volatility models. The first panel in Figure 3.15 and the fourth panel in



Table 3.7 illustrate nicely, that the mean of the jump measure can only be captured by the inclusion of a jump process into the continuous–time stochastic volatility models. Moreover, none of the models seems to be able to introduce an adequate correlation between the returns and the logarithm of realized variance, indicating that the leverage effect is not adequately modeled. The last panel of Figure 3.15 and the corresponding panel in Table 3.7 show the model–implied distribution of the Ljung–Box statistics on own serial correlation (up to order 22—corresponding roughly to one month) of the logarithmic realized variance. Both support our previous findings based on the sum of the HAR coefficients, i.e. the one–factor models tend to underestimate the persistence, whereas the two–factor model and the one–factor model with jumps tend to overestimate it. To complement our unconditional analysis we have also computed the distribution of the standard deviation of the returns, logarithmic realized variance and logarithmic Bipower variation (see last three panel of Table 3.7). As might be expected from our previous findings, the standard deviation of the returns is strongly overestimated by the two–factor model and the one–factor model with jumps. In contrast the standard deviations of the two volatility components are captured by nearly all of the models.

In summary, the estimation results and the unconditional model assessment suggest that the incorporation of a jump process into the price dynamics is important and generally leads to an improvement in model adequacy. However, for many financial applications the performance of the models conditional on the history of the price process is of primary interest, e.g. whenever return or volatility forecasts are needed. To this end, we condition on the observed price process up to four specific dates and assess the forecasting performance of the different models via simulation. In particular, our conditioning sets are selected according to the 10%– and 90%–quantiles of the continuous–time RV volatility component (which corresponds in our sample to October 1, 1992 and July 31, 2003, respectively) and of the jump measure (corresponding to January 19, 1993, and March 28, 2001, respectively), which allows us to evaluate the models at different volatility states. We then proceed as follows: conditional on the specific information sets, we compute the return, the realized variance and the Bipower variation series 22 steps ahead by simulating from the auxiliary model with parameter values implied by the stochastic volatility models, i.e. using the parameter estimates presented in Table 3.6. Repeating this procedure 10,000 times we obtain the model–implied conditional distributions of the returns, the realized variance and the Bipower variation over the next 22 days, i.e. over approximately the next month.

**Table 3.8:** Summary Statistics of Model-Implied Conditional Distributions

model	mean	std	$Q_{0.05}$	$Q_{0.25}$	$Q_{0.50}$	$Q_{0.75}$	$Q_{0.95}$
1992/10/01							
$\sum_{i=1}^{22} r_{t+i}$							
AUX	0.4781	3.2355	-4.7671	-1.2567	0.5470	2.4670	5.0891
AFF1V	0.8016	4.4195	-6.5978	-1.6377	0.8567	3.5961	7.4276
AFF1V-J-75	0.7641	2.9151	-4.1213	-0.8717	0.7767	2.6050	5.1146
AFF2V	0.4721	2.7885	-4.2930	-1.1284	0.4975	2.2908	4.6647
LL1V	0.4792	4.4486	-7.0181	-2.1454	0.4957	3.3558	7.2846
$\log(\sum_{i=1}^{22} RV_{t+i})$							
AUX	2.0632	0.4906	1.3249	1.7266	2.0029	2.3573	2.9017
AFF1V	2.7542	0.5890	1.8428	2.3570	2.6973	3.1161	3.7509
AFF1V-J-75	1.9613	0.5936	1.0172	1.5716	1.9058	2.3270	2.9561
AFF2V	1.9132	0.4732	1.1785	1.6025	1.8672	2.1929	2.7363
LL1V	2.8321	0.4917	2.0414	2.4977	2.7970	3.1561	3.6561
$\log(\sum_{i=1}^{22} BV_{t+i})$							
AUX	2.0034	0.4972	1.2521	1.6672	1.9424	2.3004	2.8515
AFF1V	2.7297	0.5934	1.8085	2.3273	2.6724	3.0957	3.7299
AFF1V-J-75	1.8965	0.5954	0.9530	1.5025	1.8405	2.2643	2.9080
AFF2V	1.8907	0.4783	1.1415	1.5760	1.8406	2.1758	2.7182
LL1V	2.8062	0.4943	2.0063	2.4715	2.7686	3.1325	3.6311
2003/07/31							
$\sum_{i=1}^{22} r_{t+i}$							
AUX	0.0791	5.4094	-8.7130	-2.8480	0.2087	3.4471	7.8329
AFF1V	0.6718	5.1995	-8.0720	-2.2933	0.6964	4.0234	8.4396
AFF1V-J-75	0.4505	5.6673	-9.0179	-2.7474	0.4632	4.0443	8.9095
AFF2V	-0.0701	5.2359	-8.9413	-3.1026	-0.0389	3.3627	7.8308
LL1V	0.3881	4.9276	-7.8418	-2.5731	0.3750	3.6246	7.9234
$\log(\sum_{i=1}^{22} RV_{t+i})$							
AUX	3.1001	0.4866	2.3647	2.7681	3.0405	3.3881	3.9320
AFF1V	3.1086	0.5525	2.2562	2.7388	3.0516	3.4436	4.0564
AFF1V-J-75	3.2971	0.5860	2.3666	2.9109	3.2403	3.6599	4.2818
AFF2V	3.1766	0.4685	2.4515	2.8679	3.1309	3.4523	3.9916
LL1V	3.0527	0.4663	2.3011	2.7351	3.0174	3.3593	3.8469
$\log(\sum_{i=1}^{22} BV_{t+i})$							
AUX	3.0371	0.4937	2.2908	2.7042	2.9759	3.3311	3.8821
AFF1V	3.0842	0.5574	2.2240	2.7085	3.0266	3.4206	4.0419
AFF1V-J-75	3.2326	0.5879	2.3040	2.8433	3.1767	3.5962	4.2307
AFF2V	3.1531	0.4739	2.4112	2.8431	3.1043	3.4342	3.9753
LL1V	3.0277	0.4691	2.2697	2.7069	2.9919	3.3366	3.8275

*Continued on next page*

model	mean	std	$Q_{0.05}$	$Q_{0.25}$	$Q_{0.50}$	$Q_{0.75}$	$Q_{0.95}$
1993/01/19							
$\sum_{i=1}^{22} r_{t+i}$							
AUX	0.3825	3.7694	-5.7272	-1.6385	0.4626	2.7005	5.7578
AFF1V	0.7497	4.7009	-7.1703	-1.8969	0.7969	3.7514	7.7874
AFF1V-J-75	0.6797	3.6280	-5.3905	-1.3501	0.6985	2.9688	6.0756
AFF2V	0.3350	3.3896	-5.4385	-1.6090	0.3635	2.5465	5.4397
LL1V	0.4387	4.6305	-7.3596	-2.2836	0.4451	3.4527	7.5127
$\log(\sum_{i=1}^{22} RV_{t+i})$							
AUX	2.3681	0.4918	1.6276	2.0311	2.3078	2.6625	3.2064
AFF1V	2.8891	0.5745	2.0024	2.5044	2.8313	3.2395	3.8632
AFF1V-J-75	2.3981	0.5944	1.4545	2.0076	2.3417	2.7648	3.3948
AFF2V	2.3018	0.4755	1.5644	1.9895	2.2547	2.5816	3.1299
LL1V	2.9185	0.4820	2.1397	2.5882	2.8852	3.2325	3.7288
$\log(\sum_{i=1}^{22} BV_{t+i})$							
AUX	2.3089	0.4984	1.5562	1.9729	2.2469	2.6050	3.1596
AFF1V	2.8646	0.5791	1.9634	2.4751	2.8063	3.2190	3.8494
AFF1V-J-75	2.3332	0.5963	1.3893	1.9381	2.2770	2.7016	3.3461
AFF2V	2.2795	0.4806	1.5283	1.9634	2.2293	2.5655	3.1120
LL1V	2.8925	0.4847	2.1081	2.5602	2.8568	3.2102	3.7037
2001/03/28							
$\sum_{i=1}^{22} r_{t+i}$							
AUX	-0.6266	7.8098	-13.3029	-4.8407	-0.4351	4.2170	10.6060
AFF1V	0.6451	5.3191	-8.2840	-2.4106	0.6617	4.0790	8.6185
AFF1V-J-75	0.0272	8.2523	-13.6895	-4.5904	0.0419	5.2179	12.3388
AFF2V	-0.8231	8.1280	-14.6242	-5.4758	-0.7471	4.4487	11.4188
LL1V	0.3742	4.9459	-7.8882	-2.6038	0.3552	3.6183	7.9319
$\log(\sum_{i=1}^{22} RV_{t+i})$							
AUX	3.8298	0.4922	3.0863	3.4943	3.7688	4.1201	4.6753
AFF1V	3.1544	0.5530	2.2991	2.7849	3.0952	3.4908	4.1049
AFF1V-J-75	4.0406	0.5957	3.0957	3.6481	3.9805	4.4077	5.0440
AFF2V	4.0449	0.4832	3.2982	3.7279	3.9964	4.3293	4.8857
LL1V	3.0581	0.4698	2.3011	2.7391	3.0235	3.3657	3.8622
$\log(\sum_{i=1}^{22} BV_{t+i})$							
AUX	3.7694	0.4992	3.0176	3.4334	3.7070	4.0640	4.6227
AFF1V	3.1295	0.5580	2.2711	2.7540	3.0720	3.4673	4.0923
AFF1V-J-75	3.9762	0.5977	3.0318	3.5803	3.9181	4.3446	4.9961
AFF2V	4.0221	0.4886	3.2590	3.7004	3.9702	4.3137	4.8693
LL1V	3.0322	0.4728	2.2674	2.7094	2.9963	3.3428	3.8358

*Continued on next page*

model	mean	std	$Q_{0.05}$	$Q_{0.25}$	$Q_{0.50}$	$Q_{0.75}$	$Q_{0.95}$
-------	------	-----	------------	------------	------------	------------	------------

Notes: The table reports the summary statistics, i.e. the mean, standard deviation and the different quantiles, of the model-implied conditional distributions of the 22-steps ahead returns, the logarithm of the 22-steps ahead realized variance and Bipower variation. The dates correspond to the conditioning set, i.e. the 10%-quantile of realized variance (first panel), the 90%-quantile of realized variance (second panel), the 10%-quantile of the jump measure (third panel), and the 90%-quantile of the jump measure (last panel). The distributions are based on a simulation study, which is further described in Section 3.5. For an explanation of the model abbreviations see Table 3.1.

Table 3.8 shows the summary statistics of these conditional distributions along with those based on the auxiliary model. The results are pretty much in line with our previous findings. In particular, conditional on low volatility or jump states the one-factor stochastic volatility models tend to overestimate the logarithmic one-month-ahead realized volatility and Bipower variation, whereas for high volatility/jump states they seem to underestimate these quantities. This behavior might be attributable to the observed tendency of these models to underestimate the volatility persistence. In contrast, the two-factor model and the one-factor model with jumps provide more accurate one-month-ahead forecasts and exhibit only a slight tendency to underestimate/overestimate the volatilities if we condition on low/high volatility states. This is also in line with the results reported earlier, i.e. that these models tend to overestimate the volatility persistence but generally produce more variation in the persistence (as can be seen from the standard deviations of the Ljung–Box statistics of these models, see e.g. Figure 3.15 (last panel)). Further, irrespective of the conditioning set all of the models provide more or less accurate forecasts of the returns. However, focusing on the tails of the return distribution reveals that, throughout, the two-factor model and the one-factor model with jumps are very accurate and outperform the one-factor models. This finding is of major importance for many risk management applications, such as Value-at-Risk computations, which require as precise as possible estimates of the tail quantiles, in particular of the left-hand tail.

## 3.6 Summary

Motivated by the inability of the existing empirical studies to provide a clear distinction between pure-diffusion models and lower-order models with jumps using daily data, we explore the information contained in the high-frequency financial data and re-assess the adequacy of the affine and logarithmic jump-diffusion models. Furthermore, our model evaluation is conducted using the general scientific modeling method of Gallant and McCulloch (2005), which generally allows the comparison of rather diverse structural models within one unifying framework. For the GSM method to be applicable, we have derived a highly accurate auxiliary model for daily returns, realized variations and jumps, summarizing the most important information inherent in the high-frequency returns. In particular, even though the discrete-time model has no direct continuous-time analog, and may in fact be consistent with many different continuous-time formulations, it is nonethe-

less highly informative about the general features that need to be accounted for in the data. We explore this fact by using its likelihood function as a summary of the data to estimate and empirically assess different continuous–time stochastic volatility models.

To this end, it turns out that the GSM method is very well suited. In particular, model evaluation is facilitated if it is conducted in terms of the auxiliary model allowing a direct comparison of the different models to capture important features of the data as given by the parameter estimates of the auxiliary model. Moreover, since not all data characteristics can be represented by individual parameters, we also perform an unconditional as well as a conditional simulation study. Note, that within the GSM method the conditional model assessment is strongly simplified when compared to other estimation methods, such as the efficient method of moments (EMM), for which the model–implied volatility states need to be extracted first, e.g. via reprojection in EMM. Instead, GSM allows to evaluate the models in terms of the auxiliary model, and as such does not require the filtering step.

Our statistical assessment of different affine and logarithmic continuous–time stochastic volatility models reveals that all of these models still miss some important features of the data, in particular they are unable to reproduce the volatility persistence as well as the leverage effect. Moreover, the inclusion of jumps into the price dynamics is an important feature that leads to an improvement in the model’s fit. Our analysis also shows, that the one–factor models underestimate the volatility persistence, whereas the two–factor model and the one–factor model with jumps tend to an overestimation but can reproduce a larger range of the persistence. Importantly, based on the conditional model evaluation we also find that both of these models provide quite accurate extreme quantile forecasts of the returns, which is essential for an adequate risk management. Overall the results suggest, that at least two factors or one factor and a jump process are needed, in order to better account for the volatility persistence and the fat–tailedness of the returns, an observation that was also made in Chernov et al. (2003) using daily data. However, other jump or stochastic volatility specifications—than the ones considered here—might be more adequate. As such it will be interesting to extend the analysis, which is straightforward using GSM regardless of the very different types of stochastic volatility models one might want to consider, as long as it is possible to simulate from them.

## 4 A Multivariate Extension of the Ornstein–Uhlenbeck Stochastic Volatility Model

A wide range of different univariate continuous–time stochastic volatility models has been developed in the financial literature aiming at capturing the most distinct features of the price process of a financial asset. As pointed out earlier, the adequacy of such models is essential for an adequate risk management and derivative pricing. For the latter application closed–form expressions are desirable imposing additional constraints on the flexibility of a stochastic volatility model. In a multivariate context, modeling becomes even more challenging. Next to capturing the individual dynamics the model also needs to reproduce the comovements and spill–over effects across different assets. In particular, knowing the correlation structure is crucial for financial decision–making, such as optimal portfolio risk management or asset allocation. In addition to those requirements, there also arise some technical problems in the multivariate setting. One is given by the necessity of a positive semidefinite covariance matrix. For stochastic volatility models this implies that the instantaneous covariance should be specified by a positive semidefinite process. Moreover, if the dimension of the return vector increases the number of parameters in the model is inflated. Hence, a parsimonious but at the same time accurate specification is needed.

Given these challenges the theoretical literature on multivariate stochastic volatility models has developed over the last few years, whereby the main focus was on discrete–time models as an alternative to the multivariate GARCH models, see e.g. Chib et al. (2006) and Harvey et al. (1994). Although the continuous–time specification is very important under the asset pricing perspective, we are aware of only a few papers considering continuous–time multivariate stochastic volatility models. Hubalek and Nicolato (2005) and Lindberg (2005) adopt a factor approach in which the volatility factors are independent and follow univariate positive Ornstein–Uhlenbeck type processes. The flexibility of these models, however, is accompanied by the difficulty to achieve identification, i.e. complicating the empirical application of these models. Moreover, closed–form expressions for the integrated covariance process—the main variable of interest for financial applications—is not available. In Gouriéroux (2006) the stochastic volatility is not driven by factors, but the full covariance matrix is specified as the sum of matrix squares of Gaussian Ornstein–Uhlenbeck processes, which are referred to as the Wishart autoregressive process.

Being the multivariate extension of a Cox–Ingersoll–Ross (CIR) process, the model shares the limitations and advantages of the univariate CIR model. Although it provides closed–form expressions for many applications, it also lacks a closed–form solution for the integrated covariance process.

We therefore introduce here a new continuous–time multivariate stochastic volatility model that is shown to meet the above mentioned requirements while providing a closed–form and very simple structure for the integrated covariance process. The general  $d$ -dimensional stochastic volatility model is given by

$$dY_t = (\mu + \Sigma_t \beta) dt + \Sigma_t^{1/2} dW_t, \quad Y_0 = 0, \quad (4.1)$$

whereby we recall the notation introduced in Section 2.1.2, i.e.  $Y$  denotes the  $d$ -dimensional logarithmic stock price process,  $\mu, \beta \in \mathbb{R}^d$  are parameters,  $(W_t)_{t \in \mathbb{R}^+}$  denotes the  $d$ -dimensional standard Brownian motion, and  $(\Sigma_t)_{t \in \mathbb{R}^+}$  is a stationary stochastic process with values in the positive semidefinite matrices  $\mathbb{S}_d^+$  being independent of  $(W_t)_{t \in \mathbb{R}^+}$ . The same representation has also been stated in e.g. Barndorff-Nielsen and Shephard (2001b), Barndorff-Nielsen et al. (2002) and Lindberg (2005). In our model the stochastic volatility process  $(\Sigma_t)_{t \in \mathbb{R}^+}$  is given by a Lévy–driven positive semidefinite Ornstein–Uhlenbeck type process which was recently introduced by Barndorff-Nielsen and Stelzer (2006), and is a multivariate extension of the positive non–Gaussian Ornstein–Uhlenbeck process used in the context of univariate stochastic volatility models in Barndorff-Nielsen and Shephard (2001b). We therefore refer to our model as the “multivariate OU–type stochastic volatility model”.

The remainder of this section is structured as follows. For illustrative purposes, we first discuss the univariate Lévy–driven Ornstein–Uhlenbeck stochastic volatility model introduced by Barndorff-Nielsen and Shephard (2001b) and its main characteristics. Extending this model to the multivariate case requires a multivariate Lévy–driven Ornstein–Uhlenbeck stochastic volatility process that is symmetric and positive semidefinite. Section 4.2 therefore provides a review of the positive semidefinite matrix process of OU–type involving the notion of a matrix subordinator, and derives some probabilistic properties that are useful for the derivation of our multivariate stochastic volatility model. Moreover, it establishes some important notation. Section 4.3 finally introduces the multivariate OU–type stochastic volatility model and its theoretical properties. The theoretical properties of our model provide the basis for different estimation methods that are presented in Section 4.4 along with an evaluation of their respective finite sample properties. Section 4.5 presents an empirical application of our model, and Section 4.6 concludes this chapter.



## 4.1 The Univariate Non–Gaussian OU–Type Stochastic Volatility Model

In contrast to the univariate stochastic volatility models considered in Section 3.1, Barndorff-Nielsen and Shephard (2001a,b, 2002a) suggest to model stochastic volatility through Lévy-driven Ornstein–Uhlenbeck-type processes rather than using Brownian motions. In particular, with an appropriate choice of the non-Gaussian Lévy process the Ornstein–Uhlenbeck type model results not only in a non-Gaussian and quite flexible law of the returns, which could capture the observed tail-thickness, but also incorporates the volatility persistence.

The stationary *Lévy-driven Ornstein–Uhlenbeck process*  $(\sigma_t^2)_{t \geq 0}$  is defined as

$$\sigma_t^2 = \sigma(0) \exp\{-\lambda t\} + \int_0^t \exp\{-\lambda(t-s)\} dz(s), \quad (4.2)$$

where  $\lambda$  is a parameter taking on only positive values, and  $(z_t)_{t \in \mathbb{R}}$  denotes a Lévy-process that is oftentimes referred to as the *Background-driving Lévy process*. Obviously, if  $z$  is a Brownian motion,  $\sigma^2$  is the Gaussian Ornstein–Uhlenbeck process, a specification that we have already considered in the logarithmic jump–diffusion models in Section 3.1. Using any other Lévy process, the resulting  $\sigma$  is a non-Gaussian Ornstein–Uhlenbeck process that is only driven by jumps. Moreover, for stochastic volatility modeling the OU process is required to be positive, unless the process is transformed adequately, e.g. by exponentiating it as in the case of the logarithmic jump–diffusion models. However, by using a positive OU process the model is linear and hence mathematically more tractable. In the non-Gaussian Ornstein–Uhlenbeck-type model of Barndorff-Nielsen and Shephard (2001b), it is therefore assumed that the BDLP satisfies the conditions of a *subordinator*, i.e. it is a pure jump Lévy process with positive increments, yielding a positive non-Gaussian OU process.

Given this is the case, then the univariate non-Gaussian OU-type stochastic volatility model is given by

$$dy_t = (\mu + \beta \sigma_t^2) dt + \sigma_t dw_t \quad (4.3)$$

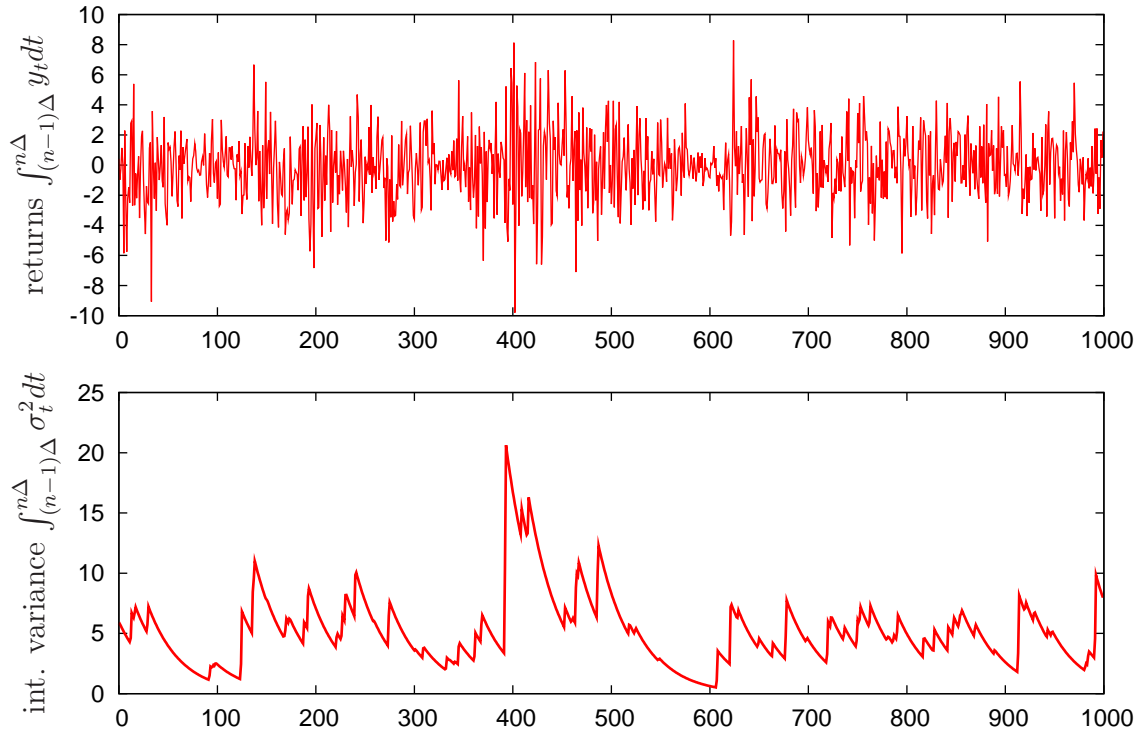
$$d\sigma_t^2 = -\lambda \sigma_t^2 dt + dz_t, \quad (4.4)$$

whereby  $\sigma_t^2$  satisfies the stochastic differential equation (4.4) and  $z$  is a subordinator without drift. The local drift of the price process is given by a constant and a risk-premium effect  $\beta \sigma_t^2$  as is also oftentimes considered in the related ARCH-in-Mean models (see e.g. the discussion in Bollerslev et al., 1988). From equation (4.4) it is obvious, that the local behavior of the integrated variance is only driven by the jumps in  $z$ , and is exponentially decaying between two jumps.<sup>1</sup> The BDLP is therefore the key ingredient of the OU-type models, since it determines the dynamics of the instantaneous (or spot) volatility and its marginal distribution. Moreover, it

---

<sup>1</sup>This interpretation of course only holds for  $z$  being of finite activity.





**Figure 4.1:** Simulated sample paths of a univariate non–Gaussian OU stochastic volatility model. The figure shows the time evolution of the simulated returns and of the integrated variance.

also determines the distribution of the returns. Note that the models can be built either by specifying a one–dimensional distribution for  $\sigma_t^2$  and by working out the implied behavior of the BDLP, or by specifying the BDLP. Barndorff-Nielsen and Shephard (2001b), for example, consider the Gamma– and the Inverse–Gaussian OU processes.

Figure 4.1 illustrates the simulated sample paths of a OU stochastic volatility model where the BDLP is specified as a compound Poisson process with exponentially distributed jumps. Note that the simulation is based on high–frequency returns, more specifically on 5 minute returns, computed for 1000 days. The figure depicts the resulting daily return and daily integrated variance series. It exemplifies that the increments of the integrated variance are only driven by jumps and exponential decay in no–jump periods. Moreover, large (small) jumps usually introduce high (low)–volatility states over several days, inducing the observed volatility clustering in the return series.

An important feature of the non–Gaussian OU–type stochastic volatility models for financial applications, is the availability of a closed–form expression for the

integrated variance, which is given by

$$\int_0^t \sigma_s^2 ds = (1/\lambda)(z_t - \sigma_t^2 - \sigma_0^2). \quad (4.5)$$

Barndorff-Nielsen and Shephard (2001b) also show that the integrated variance follows a (constrained) ARMA(1,1) process, a results that is useful for constructing a state space representation for realized variance (see Barndorff-Nielsen and Shephard, 2002a), which in turn can be used for the estimation of the model using the Kalman filter. In addition, the model allows to derive a state space representation for the squared returns leading to an alternative estimation method. Moreover, the model provides closed-form expressions for the autocovariance function of the squared returns. Although the autocovariance function is already quite flexible, in fact even more flexible, monotone decreasing forms of the volatility’s autocorrelation structure can be obtained by considering a convex combination of several Lévy-driven Ornstein–Uhlenbeck stochastic volatility factors, i.e. by a *superposition of positive OU-type processes*. Next to these appealing results, which facilitate the estimation of the model, Barndorff-Nielsen and Shephard (2002a) show that within this model option prices can be derived analytically. Further properties of the model, which, however, are shared with other stochastic volatility models, are that the returns are scaled mixtures of normals with time-dependent scaling, such that the non-Gaussianity of returns and the volatility clustering can be reproduced; and that the returns aggregate to Gaussianity as the sampling frequency increases.

Given the analytic tractability of this class of models and its empirical properties, it will be interesting to consider its multivariate extension, which will be done in the sequel. Before deriving the multivariate OU-type stochastic volatility model, however, the next section establishes some notation and introduces the multivariate counterparts of the positive OU-type processes recently proposed by Barndorff-Nielsen and Stelzer (2006) involving the concept of a matrix subordinator, along with the presentation of some important properties of these processes.

## 4.2 Positive Semidefinite Processes of OU-Type

Although an adequate transformation would allow us to use any OU-type process for modeling a positive semidefinite matrix, i.e. modeling stochastic volatility, we follow Barndorff-Nielsen and Shephard (2001a,b, 2002a) for the ease of mathematical tractability and consider only the positive semidefinite OU-type processes directly, which have recently been introduced by Barndorff-Nielsen and Stelzer (2006). In fact the authors already point out, that these processes can be used for the multivariate extension of the univariate non-Gaussian OU-type stochastic volatility model. After introducing some notation, Section 4.2.2 reviews the positive semidefinite OU-type processes and the most important results concerning the stationary distribution and the second order moment structure of the process. Moreover, we derive some further properties, e.g. we show that the sta-

tionary distribution of the positive semidefinite OU–type process can be inferred from the characteristic function of the Background–driving Lévy process—a matrix subordinator—implying that, similarly to the univariate case, the OU–type process can be constructed through the specification of the BDLP. In Section 4.2.3 we restate the result derived in Barndorff-Nielsen and Stelzer (2006) that the integrated positive semidefinite OU–type process has a simple representation, which is important for financial applications. Section 4.2.4 examines the relation of the individual components of the matrix valued OU–type process to the univariate OU–type process as considered in Barndorff-Nielsen and Shephard (2001a,b, 2002a) showing that the individual components can be represented by a superposition of univariate, but generally dependent, OU–type processes.

### 4.2.1 Notation

For the remainder of this work we write throughout  $\mathbb{R}^+$  for the positive real numbers including zero,  $\mathbb{R}^{++}$  when zero is excluded and we denote the set of real  $m \times n$  matrices by  $M_{m,n}(\mathbb{R})$ . If  $m = n$  we simply write  $M_n(\mathbb{R})$  and denote the linear subspace of symmetric matrices by  $\mathbb{S}_n$ , the (closed) positive semidefinite cone by  $\mathbb{S}_n^+$  and the open (in  $\mathbb{S}_n$ ) positive definite cone by  $\mathbb{S}_n^{++}$ .  $I_n$  stands for the  $n \times n$  identity matrix,  $\sigma(A)$  for the spectrum (the set of all eigenvalues) of a matrix  $A \in M_n(\mathbb{R})$  and  $\rho(A)$  for its spectral radius. The natural ordering on the symmetric  $n \times n$  matrices will be denoted by  $\leq$ , i.e. for  $A, B \in \mathbb{S}_n(\mathbb{R})$  we have that  $A \leq B$ , if and only if  $B - A \in \mathbb{S}_n^+$ . The Kronecker (tensor) product of two matrices  $A, B$  is written as  $A \otimes B$ .  $\text{vec}$  denotes the well-known vectorization operator that maps the  $n \times n$  matrices to  $\mathbb{R}^{n^2}$  by stacking the columns of the matrices below one another. For more information regarding the tensor product and  $\text{vec}$ –operator we refer to Horn and Johnson (1991, Chapter 4). Likewise  $\text{vech} : \mathbb{S}_d \rightarrow \mathbb{R}^{d(d+1)/2}$  denotes the “vector–half” operator that stacks the columns of the lower triangular part of a symmetric matrix below another. Finally,  $A^T$  is the transpose of a matrix  $A \in M_n(\mathbb{R})$ .

For a matrix  $A$  we denote by  $A_{ij}$  the element in the  $i$ -th row and  $j$ -th column and this notation is extended to processes in a natural way.

Regarding all random variables and processes we assume that they are defined on a given appropriate filtered probability space  $(\Omega, \mathcal{F}, P, (\mathcal{F}_t))$  satisfying the usual hypotheses. With random functions and measures we usually do not state the dependence on  $\omega \in \Omega$  explicitly.

Furthermore, we employ an intuitive notation with respect to the stochastic integration with matrix–valued integrators. Let  $A_t \in M_{m,n}(\mathbb{R})$ ,  $L_t \in M_{n,r}(\mathbb{R})$  and  $B_t \in M_{r,s}(\mathbb{R})$  be three stochastic processes then we denote by  $\int A_t dL_t B_t$  the matrix  $C$  in  $M_{m,s}(\mathbb{R})$  which has  $ij$ -th element  $C_{ij} = \sum_{k=1}^n \sum_{l=1}^r \int a_{ik} b_{lj} dL_{kl}$ . Equivalently such an integral can be understood in the sense of Métivier and Pellaumail (1980), resp. Métivier (1982), by identifying it with the integral  $\int \mathbf{A}_t dL_t$  with  $\mathbf{A}_t$  being for each fixed  $t$  the linear operator  $M_{n,r}(\mathbb{R}) \rightarrow M_{m,s}(\mathbb{R})$ ,  $X \mapsto A_t X B_t$ . Moreover, we always denote by  $\int_a^b$  with  $a \in \mathbb{R} \cup \{-\infty\}$ ,  $b \in \mathbb{R}$  the integral over the half–open

interval  $(a, b]$  for notational convenience. If  $b = \infty$  the integral is understood to be over  $(a, b)$ .

### 4.2.2 Definition and Probabilistic Properties

In the following we provide the definition of the positive semidefinite Ornstein–Uhlenbeck–type processes along with a characterization of their stationary distribution. The construction of these processes is based on a special type of matrix-valued Lévy process, i.e. the matrix subordinator, which is studied in detail in Barndorff-Nielsen and Pérez-Abreu (2007), and is defined as follows.

**Definition 4.2.1.** *An  $\mathbb{S}_d$ -valued Lévy process  $L$  is said to be a matrix subordinator, if  $L_t - L_s \in \mathbb{S}_d^+$  for all  $s, t \in \mathbb{R}^+$  with  $t > s$ .*

It can easily be shown that the paths of a matrix subordinator are  $\mathbb{S}_d^+$ -increasing and of finite variation. Moreover, the trace  $\text{tr}(L)$  is a one-dimensional (Lévy) subordinator with "tr" denoting the usual trace functional.

Given this definition the existence of Ornstein–Uhlenbeck–type processes assuming values only in the positive semidefinite matrices is ensured by the following theorem.

**Theorem 4.2.2** (Barndorff-Nielsen and Stelzer (2006, Theorem 4.5)). *Let  $(L_t)_{t \in \mathbb{R}}$  be a matrix subordinator with  $E(\log^+ \|L_t\|) < \infty$  and  $A \in M_d(\mathbb{R})$  such that  $\sigma(A) \subset (-\infty, 0) + i\mathbb{R}$ . Then the stochastic differential equation of Ornstein–Uhlenbeck type*

$$d\Sigma_t = (A\Sigma_t + \Sigma_t A^T)dt + dL_t$$

*has a unique stationary solution*

$$\Sigma_t = \int_{-\infty}^t e^{A(t-s)} dL_s e^{A^T(t-s)}$$

*or, in vectorial representation,*

$$\text{vec}(\Sigma_t) = \int_{-\infty}^t e^{(I_d \otimes A + A \otimes I_d)(t-s)} d\text{vec}(L_s).$$

*Moreover,  $\Sigma_t \in \mathbb{S}_d^+$  for all  $t \in \mathbb{R}$ .*

Note, that Stelzer (2006) has recently shown that the linear operator  $\mathbf{A} : \mathbb{S}_d \rightarrow \mathbb{S}_d$  given by  $X \mapsto AX + XA^T$  with  $A \in M_d(\mathbb{R})$  uniquely identifies  $A$ .<sup>2</sup> Moreover, Stelzer (2006) establishes that all linear operators  $\mathbf{A}$  which satisfy  $\exp(\mathbf{A}t)(\mathbb{S}_d) = \mathbb{S}_d^+ (*)$  for all  $t \in \mathbb{R}$  are necessarily of the form  $X \mapsto AX + XA^T$  for some  $A \in M_d(\mathbb{R})$ .

<sup>2</sup>In particular, to identify  $A$  it is already sufficient to know the values of  $\mathbf{A}E_{ii}$  for  $i = 1, \dots, d$  where  $E_{ii}$  are the  $d \times d$  matrices with only zero entries except for one entry of one at the  $i$ -th diagonal element.

The property (\*) is important for the well-definedness of positive semidefinite OU-type processes. The equality also immediately implies that the distribution of the OU-type process  $\Sigma_t$  is not concentrated on any proper linear subspace of  $\mathbb{S}_d$  when the distribution of the driving Lévy process  $L_t$  is not concentrated on any proper linear subspace.

Having defined the positive semidefinite OU-type process, we now turn to the characterization of its stationary distribution. To this end, note that Barndorff-Nielsen and Pérez-Abreu (2007) show that the subordinator  $L_t$  has the following characteristic function

$$\mu_{L_t}(Z) = \exp\left(t\psi_L(Z)\right), \quad Z \in \mathbb{S}_d, \quad \text{where} \quad (4.6)$$

$$\psi_L(Z) := i\text{tr}(\gamma_L Z) + \int_{\mathbb{S}_d^+ \setminus \{0\}} (e^{i\text{tr}(XZ)} - 1)\nu_L(dX). \quad (4.7)$$

Based on this characteristic function, the stationary distribution of the positive semidefinite Ornstein–Uhlenbeck process and its second-order-moment structure, which are important for the derivation of the second order structure of the stochastic volatility model, are then given as follows.

**Proposition 4.2.3** (Barndorff-Nielsen and Stelzer (2006, Proposition 4.7)). *The stationary distribution of the positive semidefinite Ornstein–Uhlenbeck process  $\Sigma_t$  is infinitely divisible with characteristic function*

$$\begin{aligned} \hat{\mu}_\Sigma(Z) &= \exp\left(\int_0^\infty \psi_L\left(e^{A^T s} Z e^{A s}\right) ds\right) \\ &= \exp\left(i\text{tr}(\gamma_\Sigma Z) + \int_{\mathbb{S}_d^+ \setminus \{0\}} (e^{i\text{tr}(XZ)} - 1)\nu_\Sigma(dX)\right), \quad Z \in \mathbb{S}_d, \end{aligned} \quad (4.8)$$

where

$$\gamma_\Sigma = -\mathbf{A}^{-1}\gamma_L$$

with  $\mathbf{A}$  defined as the linear operator  $\mathbf{A} : M_d(\mathbb{R}) \rightarrow M_d(\mathbb{R})$ ,  $X \mapsto AX + XA^T$  which can be represented as  $\text{vec}^{-1} \circ ((I_d \otimes A) + (A \otimes I_d))^{-1} \circ \text{vec}$  and

$$\nu_\Sigma(E) = \int_0^\infty \int_{\mathbb{S}_d^+ \setminus \{0\}} I_E(e^{A s} x e^{A^T s}) \nu_L(dx) ds$$

for all Borel sets  $E$  in  $\mathbb{S}_d^+ \setminus \{0\}$ .

Assume that the driving Lévy process is square-integrable. Then the second order moment structure is given by

$$E(\Sigma_t) = \gamma_\Sigma - \mathbf{A}^{-1} \int_{\mathbb{S}_d^+ \setminus \{0\}} y \nu_L(dy) = -\mathbf{A}^{-1} E(L_1) \quad (4.9)$$

$$\begin{aligned} \text{var}(\text{vec}(\Sigma_t)) &= \int_0^\infty e^{(A \otimes I_d + I_d \otimes A)t} \text{var}(\text{vec}(L_1)) e^{(A^T \otimes I_d + I_d \otimes A^T)t} dt \\ &= -\mathcal{A}^{-1} \text{var}(\text{vec}(L_1)) \end{aligned} \quad (4.10)$$

$$\text{cov}(\text{vec}(\Sigma_{t+h}), \text{vec}(\Sigma_t)) = e^{(A \otimes I_d + I_d \otimes A)h} \text{var}(\text{vec}(\Sigma_t)), \quad (4.11)$$

where  $t \in \mathbb{R}$  and  $h \in \mathbb{R}^+$  and  $\mathcal{A} : M_{d^2}(\mathbb{R}) \rightarrow M_{d^2}(\mathbb{R})$ ,  $X \mapsto (A \otimes I_d + I_d \otimes A)X + X(A^T \otimes I_d + I_d \otimes A^T)$ . The linear operator  $\mathcal{A}$  can be represented as

$$\text{vec}^{-1} \circ ((I_{d^2} \otimes (A \otimes I_d + I_d \otimes A)) + ((A \otimes I_d + I_d \otimes A) \otimes I_{d^2})) \circ \text{vec}.$$

Note, that equation (4.10) can alternatively be restated as

$$-\text{var}(\text{vec}(L_1)) = (A \otimes I_d + I_d \otimes A)\text{var}(\text{vec}(\Sigma_t)) + \text{var}(\text{vec}(\Sigma_t))(A^T \otimes I_d + I_d \otimes A^T).$$

Moreover, we used the  $\text{vec}$ -operator above, as this clarifies the order of the elements of the (co)variance matrix. One might wonder why one does not use the “vector half” operator  $\text{vech}$  that stacks the columns of the lower diagonal part (including the diagonal) of a symmetric matrix below one another. Although this would avoid the redundancies in the covariance matrices  $\text{var}(\text{vec}(L_1))$  and  $\text{var}(\text{vec}(\Sigma_t))$  caused by the symmetry of  $L_1$  and  $\Sigma_t$ , it seems to be rather disadvantageous when seeking explicit expressions, since to the best of our knowledge, there are far less results from linear algebra available for the  $\text{vech}$ -operator than for the  $\text{vec}$ -operator. Hence, we will use the  $\text{vec}$ -operator throughout most of the rest of this work and merely note that one can, of course, switch to the  $\text{vech}$ -operator by simply picking the relevant components.

Having derived the stationary distribution of the positive semidefinite OU-type process, it is important to note, that generally the finiteness of its moments is completely characterized by the driving Lévy process.

**Proposition 4.2.4.** *Let  $(\Sigma_t)_{t \in \mathbb{R}}$  be a strictly stationary OU-type process in  $\mathbb{S}_d^+$  with driving matrix subordinator  $L_t$  and be  $r \in \mathbb{R}^{++}$ . Then  $E(\|\Sigma_0\|^r) < \infty$ , if and only if  $E(\|L_1\|^r) < \infty$  or equivalently  $\int_{\mathbb{S}_d^+, \|x\| \geq 1} \|x\|^r \nu_L(dx) < \infty$ .*

*Proof.* Follows by a straightforward adaptation of the proof of Marquardt and Stelzer (2007, Proposition 3.30) to the matrix case.  $\square$

Furthermore, it is noteworthy that these processes exhibit a very nice dependence structure. To illustrate this, let us thus introduce the notion of  $\beta$ -mixing:

**Definition 4.2.5** (cf. Davydov (1973)). *A continuous-time stationary stochastic process  $X = \{X_t\}_{t \in \mathbb{R}}$  is called strongly (or  $\alpha$ -) mixing, if*

$$\alpha_l := \sup \{ |P(A \cap B) - P(A)P(B)| : A \in \mathcal{F}_{-\infty}^0, B \in \mathcal{F}_l^\infty \} \rightarrow 0$$

as  $l \rightarrow \infty$ , where  $\mathcal{F}_{-\infty}^0 := \sigma(\{X_t\}_{t \leq 0})$  and  $\mathcal{F}_l^\infty = \sigma(\{X_t\}_{t \geq l})$ .

*It is said to be  $\beta$ -mixing (or completely regular), if*

$$\beta_l := E \left( \sup \{ |P(B|\mathcal{F}_{-\infty}^0) - P(B)| : B \in \mathcal{F}_l^\infty \} \right) \rightarrow 0$$

as  $l \rightarrow \infty$ .

Note that  $\alpha_l \leq \beta_l$  and thus any  $\beta$ -mixing process is strongly mixing, which implies that many results regarding statistics can be applied.

**Proposition 4.2.6.** *Let  $\Sigma_t$  be a stationary OU-type process in  $\mathbb{S}_d^+$ . Then  $\Sigma_t$  is a temporally homogeneous strong Markov process.*

*If the driving Lévy process  $L$  satisfies, moreover,*

$$\int_{\mathbb{S}_d^+, \|x\| \geq 1} \|x\|^r \nu_L(dx) < \infty \quad (4.12)$$

*for some  $r > 0$ , then  $\Sigma_t$  is  $\beta$ -mixing with mixing coefficients  $\beta_l = O(e^{-al})$  for some  $a > 0$ . In particular,  $\Sigma_t$  is strongly mixing and ergodic.*

*Proof.* Follows from Protter (2004, Theorem V.32) and Masuda (2004, Theorem 4.3).  $\square$

Furthermore, we can obtain the conditions ensuring that the stationary OU-type process  $\Sigma_t$  is almost surely strictly positive definite.

**Theorem 4.2.7** (Barndorff-Nielsen and Stelzer (2006, Theorem 4.9)). *If  $\gamma_L \in \mathbb{S}_d^{++}$  or  $\nu_L(\mathbb{S}_d^{++}) > 0$ , then the stationary distribution  $P_\Sigma$  of  $\Sigma_t$  is concentrated on  $\mathbb{S}_d^{++}$ , i.e.  $P_\Sigma(\mathbb{S}_d^{++}) = 1$ .*

Additionally, the stationary distributions of multivariate Ornstein–Uhlenbeck-type processes are operator self-decomposable as is shown in Pigorsch and Stelzer (2007).

### 4.2.3 The Integrated Process

An important feature of the positive semidefinite Ornstein–Uhlenbeck-type processes is that the integrated process, denoted by  $\Sigma_t^+$ , has a simple representation.

**Proposition 4.2.8** (Barndorff-Nielsen and Stelzer (2006, Proposition 4.10)). *Let  $\Sigma_t$  be a positive semidefinite Ornstein–Uhlenbeck process with initial value  $\Sigma_0 \in \mathbb{S}_d^+$  and driving Lévy process  $L_t$ . Then the integrated Ornstein–Uhlenbeck process  $\Sigma_t^+$  is given by*

$$\Sigma_t^+ := \int_0^t \Sigma_s ds = \mathbf{A}^{-1} (\Sigma_t - \Sigma_0 - L_t)$$

*for  $t \in \mathbb{R}^+$ , where  $\mathbf{A}$  is the linear operator defined in Proposition 4.2.3*

The similarity of this expression to the representation of the integrated process in the univariate case, see equation 4.5 is obvious. From a financial point of view this is very appealing since the integrated process corresponds to the integrated covolatility, which is the main variable of interest in financial applications. Hence, a simple representation is therefore desirable.



#### 4.2.4 Marginal Dynamics

In this section we compare the behavior of the individual components  $\Sigma_{ij,t}$  implied by a positive semidefinite OU–type process  $\Sigma_t = (\Sigma_{ij,t})_{1 \leq i,j \leq d}$  in  $\mathbb{S}_d^+$  with the univariate OU–type processes. We therefore derive the marginal dynamics. Assume that  $A$  is real diagonalizable and  $\sigma(A) = \{\lambda_1, \dots, \lambda_d\}$ . Let  $U \in M_d(\mathbb{R})$  be such that

$$UAU^{-1} = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \lambda_d \end{pmatrix} := D. \quad (4.13)$$

Denoting  $(U^T)^{-1} = (U^{-1})^T$  by  $U^{-T}$ , we have that

$$\begin{aligned} \Sigma_t &= U^{-1} \left( \int_{-\infty}^t e^{D(t-s)} U dL_s U^T e^{D^T(t-s)} \right) U^{-T} \\ &= U^{-1} \left( \int_{-\infty}^t e^{D(t-s)} d(U L_s U^T) e^{D^T(t-s)} \right) U^{-T}. \end{aligned} \quad (4.14)$$

Defining  $M_t = U L_t U^T$  for  $t \in \mathbb{R}$  we see that  $M$  is again a Lévy process in  $M_d(\mathbb{R})$  and even a matrix subordinator. Moreover, one obtains that

$$\left( \int_{-\infty}^t e^{D(t-s)} d(U L_s U^T) e^{D^T(t-s)} \right)_{ij} = \int_{-\infty}^t e^{(\lambda_i + \lambda_j)(t-s)} dM_{ij,s},$$

which obviously shows that the individual components of

$$U \Sigma_t U^T = \int_{-\infty}^t e^{D(t-s)} d(U L_s U^T) e^{D^T(t-s)}$$

are stationary one–dimensional Ornstein–Uhlenbeck–type processes with associated SDE  $d(U \Sigma_t U^T)_{ij} = (\lambda_i + \lambda_j)(U \Sigma_t U^T)_{ij} dt + dM_{ij,t}$  (\*). Note further, that  $M_{ii}$  for  $1 \leq i \leq d$  is necessarily a subordinator and  $(U \Sigma_t U^T)_{ii}$  has to be a positive OU–type process. These assertions do, however, fail in general for  $M_{ij}$  and  $(U \Sigma_t U^T)_{ij}$  with  $i \neq j$ .

Together with (4.14) the above considerations show that the individual components  $\Sigma_{ij,t}$  of  $\Sigma_t$  are superpositions of (at most  $d^2$ ) univariate OU–type processes. However, unlike in Barndorff-Nielsen and Shephard (2001b), the individual OU processes superimposed are in general not independent.

With the obvious modifications the above results hold also true for general diagonalizable (non–real!)  $A \in M_d(\mathbb{R})$ . Then  $X^T$  denotes the Hermitian of a matrix  $X \in M_d(\mathbb{C})$  and  $U \Sigma_t U^T$  are OU–type processes in the positive semidefinite complex matrices. Note that  $(U \Sigma_t U^T)_{ii}$  still have to be real (even positive) and  $M_{ii}$  a real matrix subordinator. Furthermore, (\*) becomes  $d(U \Sigma_t U^T)_{ij} = (\lambda_i + \bar{\lambda}_j)(U \Sigma_t U^T)_{ij} dt + dM_{ij,t}$ .



### 4.3 The Multivariate OU–Type Stochastic Volatility Model

Having discussed the positive semidefinite OU–type process we can now introduce our multivariate stochastic volatility model. Recall that the general  $d$ -dimensional stochastic volatility model is given by

$$dY_t = (\mu + \Sigma_t \beta) dt + \Sigma_t^{1/2} dW_t, \quad Y_0 = 0,$$

where  $Y$  denotes the  $d$ -dimensional logarithmic price process,  $\mu, \beta \in \mathbb{R}^d$  are the drift and so-called risk premium parameters, respectively,  $(W_t)_{t \in \mathbb{R}^+}$  denotes a  $d$ -dimensional standard Brownian motion and  $(\Sigma_t)_{t \in \mathbb{R}^+}$  is a stationary stochastic process with values in  $\mathbb{S}_d^+$  and is independent of  $(W_t)_{t \in \mathbb{R}^+}$ . Moreover,  $(\Sigma_t)_{t \in \mathbb{R}^+}$  is referred to as the stochastic volatility process.

The above model has been stated in several papers (Barndorff-Nielsen and Shephard, 2001b; Barndorff-Nielsen et al., 2002; Lindberg, 2005, e.g.) along with various concrete specifications for the volatility process  $\Sigma_t$  (mainly factor models).

In this work we mainly focus on a specification in which the volatility process is given by a Lévy–driven positive semidefinite OU–type process where the driving Lévy process  $(L_t)_{t \in \mathbb{R}^+}$  and the the Brownian Motion of the price process are independent. However, whenever possible we state our result for the general model given in (4.1). Note, that for the multivariate OU–type stochastic volatility model the corresponding formula turn out to be very explicit.

Generally we follow Barndorff-Nielsen and Shephard (2001b) and presume  $Y_0 = 0$ , which is no real constraint as it just corresponds to a normalization of the prices at time zero. In the Ornstein–Uhlenbeck–type stochastic volatility model, however, we again extend the driving Lévy process to one defined on the whole real line and thus the stochastic volatility process of our multivariate Lévy–driven positive semidefinite OU–type stochastic volatility model is given by

$$\Sigma_t = \int_{-\infty}^t e^{A(t-s)} dL_s e^{A^T(t-s)}. \quad (4.15)$$

Note that this corresponds to starting the OU–type process at time zero with  $\Sigma_0$  having stationary distribution and being independent of  $(L_t)_{t \in \mathbb{R}^+}$ .

Before deriving the most important theoretical properties of the multivariate OU–type stochastic volatility model, we shortly reconsider the price and volatility increments and provide an illustration of our model. In particular, the returns over a unit time interval, such as one day, as defined in equation (2.5) are given here by

$$\mathbf{Y}_n = \int_{(n-1)\Delta}^{n\Delta} (\mu + \Sigma_t \beta) dt + \int_{(n-1)\Delta}^{n\Delta} \Sigma_t^{1/2} dW_t,$$

whereas the integrated volatility over this interval is defined by

$$\Sigma_n := \int_{(n-1)\Delta}^{n\Delta} \Sigma_t dt = \Sigma_{n\Delta}^+ - \Sigma_{(n-1)\Delta}^+.$$

As such, it follows that the returns are a scaled mixture of multivariate Normals, i.e.

$$\mathbf{Y}_n | \boldsymbol{\Sigma}_n \sim N_d(\mu\Delta + \boldsymbol{\Sigma}_n\beta, \boldsymbol{\Sigma}_n) \quad (4.16)$$

with  $N_d(m, \tau)$  denoting the  $d$ -dimensional Normal distribution with mean  $m$  and variance  $\tau$ .

To illustrate our multivariate positive semidefinite OU-type stochastic volatility model, we simulate from a bivariate model with finite active BDLP. In particular, we set  $\mu = \beta = 0$ , and assume that the BDLP is a vector compound Poisson process, i.e. has zero non-diagonal elements, with exponentially distributed jumps of the two BDLP components occurring at the same time. Figure 4.2 presents the resulting return and integrated variance series of the two assets. Obviously, our model introduces volatility clustering in both assets whereby high and low volatility periods occur roughly simultaneously. Moreover, the time evolution of the integrated variance series shows that in contrast to the univariate case (see Figure 4.1) the individual series are not anymore solely driven by jumps and an exponential decay between two consecutive jumps, but also exhibit smooth increases, which is due to the joint modeling. The comovements between the two assets is further illustrated in Figure 4.3 depicting the covariance and correlation series. As can be seen, the assets of our simulation are indeed mostly positively correlated, which is also obvious from the scatter plot of the two returns series.

Note that, it is easy to incorporate a leverage effect into the OU-type stochastic volatility model analogously to the univariate case by specifying

$$dY_t = (\mu + \Sigma_t\beta)dt + \Sigma_t^{1/2}dW_t + \psi dL_t. \quad (4.17)$$

with  $\psi$  being a linear operator from  $\mathbb{S}_d$  to  $\mathbb{R}^d$ . However, even in the univariate case it is hard to obtain any closed form expressions for the model with leverage effect, so we focus here on the model without leverage effect.

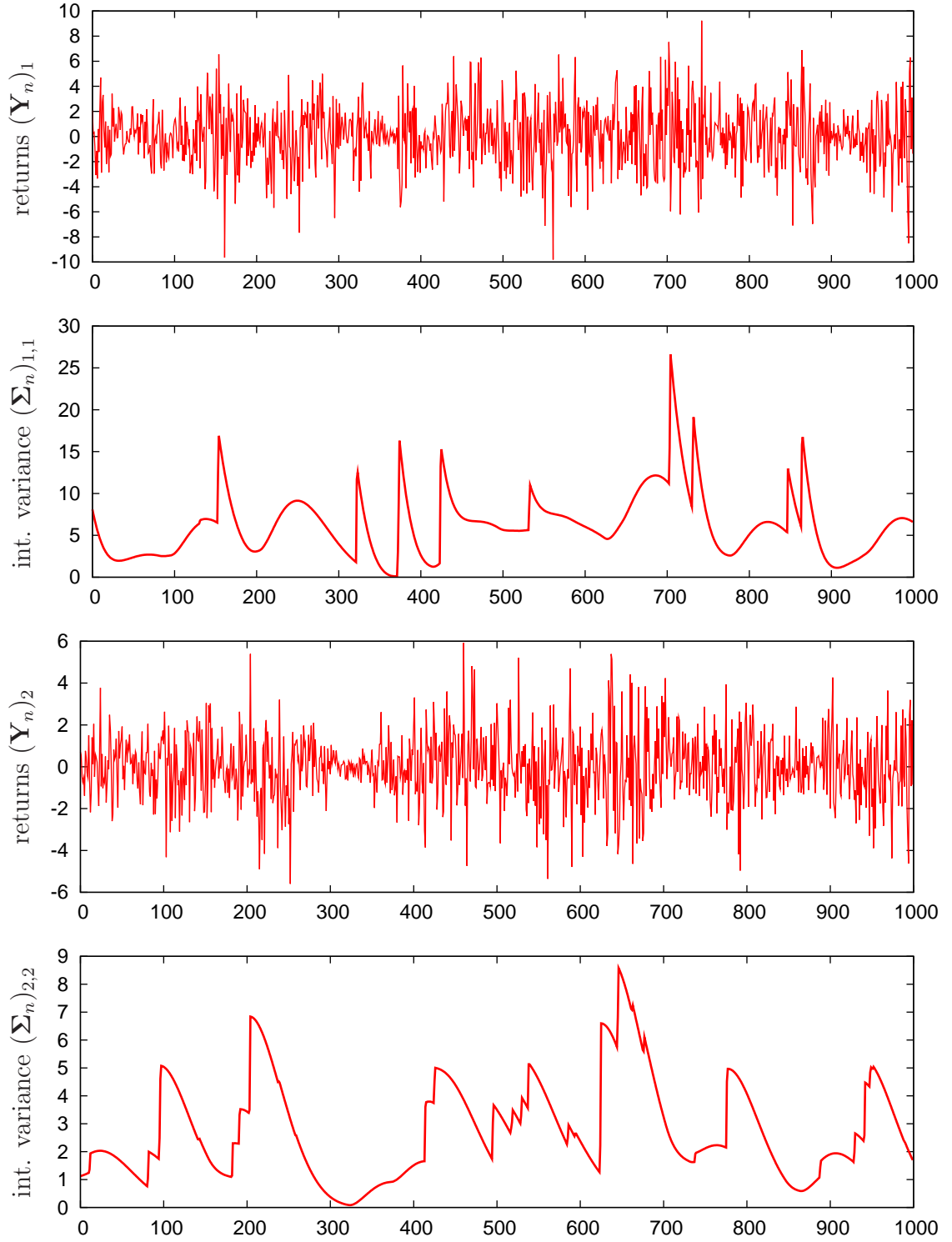
The remaining section is structured as follows. The next subsection presents the second order properties of the model. In Section 4.3.2 we introduce a state space representation for the vector of returns and squared returns, whereas section 4.3.3 derives such a representation for the realized covariation.

### 4.3.1 Second Order Structure

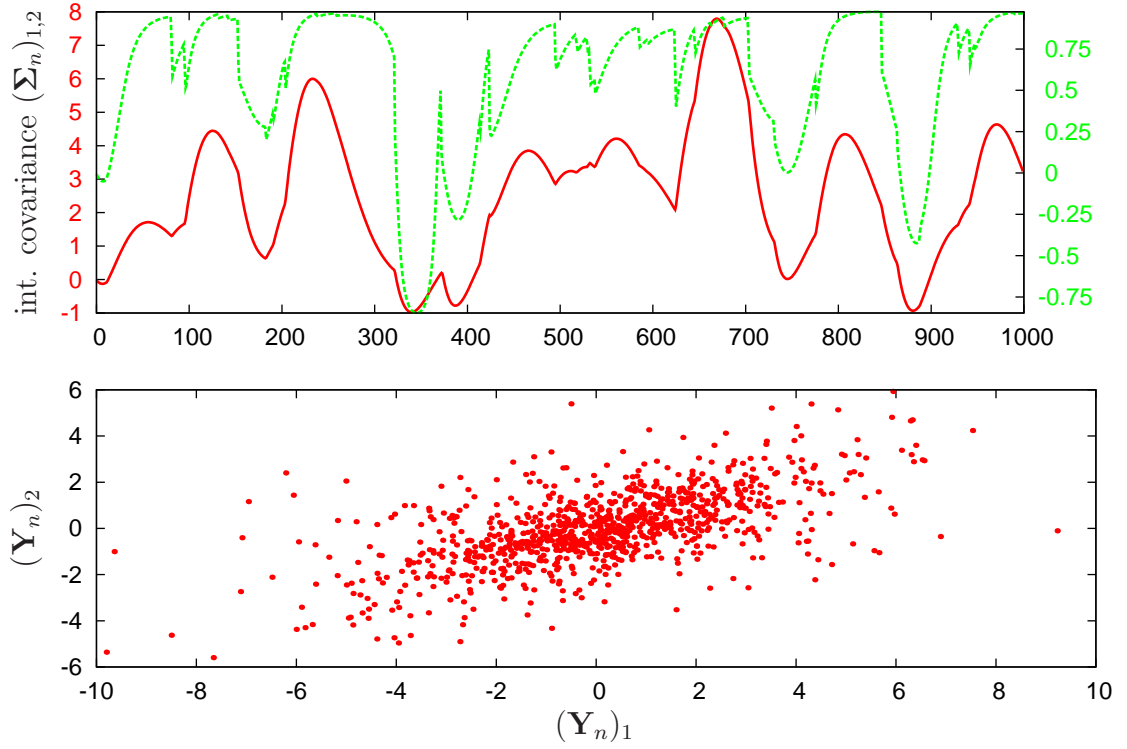
The aim of this section is to study the moments, in particular the second order structure, of various processes such as the integrated volatility and its increments, the price process, the returns and squared returns of our multi-dimensional stochastic volatility model. We consider the general model whenever possible, but the most explicit results are obtained for the OU-type stochastic volatility model. This analysis provides a basis for estimation and forecasting of our model. Henceforth, we make the assumption, that the stationary stochastic volatility process  $(\Sigma_t)_{t \in \mathbb{R}^+}$  has a finite second moment.<sup>3</sup>

---

<sup>3</sup>In the OU-case this means that the driving Lévy process has to be square-integrable.



**Figure 4.2:** Simulated sample paths of a bivariate positive semidefinite OU-type stochastic volatility model. The figure shows the simulated return series as well as the integrated variance series.



**Figure 4.3:** Simulated sample paths and return scatter plot of a bivariate positive semidefinite OU-type stochastic volatility model. The figure presents the time evolution of the integrated covariance between the two assets (red solid line, values are according to the left-hand axis) and of their correlation (green dashed line, values are according to the right-hand axis). The second panel shows the pairwise scatter plot between the returns of the two assets.

As usual we define the autocovariance function  $\text{acov}_X(h)$  by

$$\text{acov}_X(h) = \text{cov}(X_h, X_0) = E(X_h X_0^T) - E(X_0)E(X_0)^T \quad h > 0 \quad (4.18)$$

and if  $X$  is a second order stationary process in  $M_d(\mathbb{R})$  then we set  $\text{acov}_X := \text{acov}_{\text{vec}(X)}$ .

The twice integrated autocovariance function of the stationary volatility process  $\Sigma_t$  will be of particular importance. Thus we define

$$r^+(t) = \int_0^t \text{acov}_\Sigma(u) du \quad \text{and} \quad (4.19)$$

$$r^{++}(t) = \int_0^t r^+(u) du. \quad (4.20)$$

Based on these definitions we are able to show the following properties for the general multivariate stochastic volatility model.

**Theorem 4.3.1.** *For the general stochastic volatility model with  $(\Sigma_t)_{t \in \mathbb{R}^+}$  being stationary and square-integrable it holds that:*

$$E(\Sigma_t^+) = tE(\Sigma_0) \quad (4.21)$$

$$\text{var}(\text{vec}(\Sigma_t^+)) = r^{++}(t) + (r^{++}(t))^T \quad (4.22)$$

$$E(Y_t) = (\mu + E(\Sigma_0)\beta)t \quad (4.23)$$

$$\text{var}(Y_t) = E(\Sigma_0)t + (\beta^T \otimes I_d)\text{var}(\text{vec}(\Sigma_t^+))(\beta \otimes I_d). \quad (4.24)$$

Furthermore the discretely observed integrated volatility  $(\Sigma_n)_{n \in \mathbb{N}}$  is stationary and square-integrable. We have

$$E(\Sigma_n) = \Delta E(\Sigma_0) \quad (4.25)$$

$$\text{var}(\text{vec}(\Sigma_n)) = r^{++}(\Delta) + (r^{++}(\Delta))^T \quad (4.26)$$

$$\text{acov}_\Sigma(h) = r^{++}(h\Delta + \Delta) - 2r^{++}(h\Delta) + r^{++}(h\Delta - \Delta), \quad h \in \mathbb{N}. \quad (4.27)$$

Likewise we have that the discretely observed log-price increments  $(Y_n)_{n \in \mathbb{N}}$  are stationary and square-integrable. It holds that

$$E(Y_n) = (\mu + E(\Sigma_0)\beta)\Delta \quad (4.28)$$

$$\text{var}(Y_n) = E(\Sigma_0)\Delta + (\beta^T \otimes I_d)\text{var}(\text{vec}(\Sigma_\Delta^+))(\beta \otimes I_d) \quad (4.29)$$

$$\text{acov}_Y(h) = (\beta^T \otimes I_d)\text{acov}_\Sigma(h)(\beta \otimes I_d), \quad h \in \mathbb{N} \quad (4.30)$$

Let  $\Sigma_t$  now be a positive semidefinite OU-type process with driving matrix subordinator  $L$  then

$$E(\Sigma_0) = -\mathbf{A}^{-1}E(L_1) \quad (4.31)$$

$$\begin{aligned} r^{++}(t) &= (\mathcal{A}^{-2}(e^{\mathcal{A}t} - I_{d^2}) - \mathcal{A}^{-1}t) \text{var}(\text{vec}(\Sigma_0)) \\ &= -(\mathcal{A}^{-2}(e^{\mathcal{A}t} - I_{d^2}) - \mathcal{A}^{-1}t) \mathcal{A}^{-1} \text{var}(\text{vec}(L_1)) \end{aligned} \quad (4.32)$$

$$\begin{aligned} \text{acov}_\Sigma(h) &= e^{\mathcal{A}\Delta(h-1)} \mathcal{A}^{-2} (I_{d^2} - e^{\mathcal{A}\Delta})^2 \text{var}(\text{vec}(\Sigma_0)) \\ &= -e^{\mathcal{A}\Delta(h-1)} \mathcal{A}^{-2} (I_{d^2} - e^{\mathcal{A}\Delta})^2 \mathcal{A}^{-1} \text{var}(\text{vec}(L_1)), \quad h \in \mathbb{N} \end{aligned} \quad (4.33)$$

where  $\mathcal{A} := A \otimes I_d + I_d \otimes A$ .

*Proof.* The proof of Theorem 4.3.1 is given in Pigorsch and Stelzer (2007).  $\square$

Pigorsch and Stelzer (2007) show that the integrated covariance matrix  $\Sigma_n$  has a VARMA(1,1) structure. If  $\Sigma_n$  would be observable, then this result could be used for the estimation and inference of our multivariate OU-type stochastic volatility model. Instead, we can only observe the log-price increments  $\mathbf{Y}_n$ , whose second order structure, however, obviously does not allow for an in-depth analysis of the latent stochastic volatility model. The next result shows that the squared log-price increments  $\mathbf{Y}_n \mathbf{Y}_n^T$  are much more informative and can be used for model assessment. This is not surprising, as their sum converges in probability to the integrated volatility as the sampling frequency goes to infinite.

**Theorem 4.3.2.** *In the general stochastic volatility model with  $\mu = \beta = 0$  the second order structure of the squared log price process is given by*

$$E(Y_t Y_t^T) = \text{var}(Y_t) + E(Y_t)E(Y_t^T) = E(\Sigma_0)t \quad (4.34)$$

$$\begin{aligned} \text{var}(\text{vec}(Y_t Y_t^T)) &= (I_{d^2} + \mathbf{Q} + \mathbf{PQ}) (r^{++}(t) + (r^{++}(t))^T) \\ &\quad + (I_{d^2} + \mathbf{P}) (E(\Sigma_0) \otimes E(\Sigma_0)) t^2 \end{aligned} \quad (4.35)$$

and the one of the squared log return increments  $(\mathbf{Y}_n \mathbf{Y}_n^T)_{n \in \mathbb{N}}$  is given by

$$E(\mathbf{Y}_n \mathbf{Y}_n^T) = \text{var}(\mathbf{Y}_n) + E(\mathbf{Y}_n)E(\mathbf{Y}_n^T) = E(\Sigma_0)\Delta \quad (4.36)$$

$$\begin{aligned} \text{var}(\text{vec}(\mathbf{Y}_n \mathbf{Y}_n^T)) &= (I_{d^2} + \mathbf{Q} + \mathbf{PQ}) (r^{++}(\Delta) + (r^{++}(\Delta))^T) \\ &\quad + (I_{d^2} + \mathbf{P}) (E(\Sigma_0) \otimes E(\Sigma_0)) \Delta^2 \end{aligned} \quad (4.37)$$

$$\text{acov}_{\mathbf{Y} \mathbf{Y}^T}(h) = \text{acov}_{\Sigma}(h) \text{ for } h \in \mathbb{N} \quad (4.38)$$

where

$$\mathbf{P} : M_{d^2} \rightarrow M_{d^2}, \quad (4.39)$$

$$(\mathbf{P}X)_{i,(p-1)d+q} = X_{i,(q-1)d+p} \text{ for all } i = \{1, 2, \dots, d^2\}, p, q = \{1, 2, \dots, d\}$$

$$\mathbf{Q} : M_{d^2} \rightarrow M_{d^2}, \quad (4.40)$$

$$(\mathbf{Q}X)_{(k-1)d+l,(p-1)d+q} = X_{(k-1)d+p,(l-1)d+q} \text{ for all } k, l, p, q = \{1, 2, \dots, d\}$$

are linear operators.

Componentwise we have for the variance

$$\begin{aligned} \text{cov}((\mathbf{Y}_{i,n} \mathbf{Y}_{j,n}^T), (\mathbf{Y}_{k,n} \mathbf{Y}_{l,n}^T)) &= \text{var}(\text{vec}(\mathbf{Y}_n \mathbf{Y}_n^T))_{(j-1)d+i, (l-1)d+k} \\ &= \int_0^t \int_0^z (\text{cov}(\Sigma_{ij,z}, \Sigma_{kl,u}) + \text{cov}(\Sigma_{ij,u}, \Sigma_{kl,z})) \, dudz \\ &\quad + \int_0^t \int_0^z (E(\Sigma_{jl,z} \Sigma_{ik,u}) + E(\Sigma_{jl,u} \Sigma_{ik,z}) + E(\Sigma_{jk,z} \Sigma_{il,u}) + E(\Sigma_{jk,u} \Sigma_{il,z})) \, dudz \end{aligned} \quad (4.41)$$

Since the process  $(\text{vec}(\mathbf{Y}_n \mathbf{Y}_n^T))_{n \in \mathbb{N}}$  has the same autocovariance function as the integrated variance it is again a VARMA(1,1) process.

*Proof.* The proof of Theorem 4.3.2 is given in Pigorsch and Stelzer (2007).  $\square$

Given the second order properties of  $\mathbf{Y}\mathbf{Y}^T$  we are now able to estimate our model via the autocovariance function. We apply this procedure in our empirical application. Although the implementation of this estimation method is straight forward its theoretical, i.e. distributional properties are unavailable. This is also highlighted by a Monte–Carlo study, which in fact shows that the autocovariance estimator is biased for a finite number of observations. Another shortcoming is the unavailability of the underlying states of the time–varying volatility matrix. To overcome this we could apply a particle filter. Alternatively, the Kalman filter can be used for the quasi maximum likelihood estimation and volatility filtering based on a state space representation for the squared high–frequency returns, which is derived in the sequel.

### 4.3.2 State Space Representation

In this section we establish a state space representation for the joint process  $(\mathbf{Y}_n, \mathbf{Y}_n \mathbf{Y}_n^T)_{n \in \mathbb{N}}$ . This allows inference using basic tools, e.g. estimating the latent stochastic volatility via the Kalman filter. Throughout we assume  $\beta = 0$ . We first analyze the general stochastic volatility model and thereafter focus on the OU–type model

Recall first that

$$\mathbf{Y}_n = \Delta\mu + \int_{(n-1)\Delta}^{n\Delta} \Sigma_s^{1/2} dW_s,$$

which immediately implies

$$\begin{aligned} \mathbf{Y}_n \mathbf{Y}_n^T &= \Delta^2 \mu \mu^T + \int_{(n-1)\Delta}^{n\Delta} \Sigma_s^{1/2} dW_s \int_{(n-1)\Delta}^{n\Delta} dW_s^T \Sigma_s^{1/2} \\ &\quad + \Delta \int_{(n-1)\Delta}^{n\Delta} \Sigma_s^{1/2} dW_s \mu^T + \Delta \mu \int_{(n-1)\Delta}^{n\Delta} dW_s^T \Sigma_s^{1/2}. \end{aligned} \quad (4.42)$$

Defining

$$u_n := \begin{pmatrix} u_{1,n} \\ u_{2,n} \end{pmatrix} \quad (4.43)$$

$$u_{1,n} := \int_{(n-1)\Delta}^{n\Delta} \Sigma_s^{1/2} dW_s \quad (4.44)$$

$$\begin{aligned} u_{2,n} &:= \int_{(n-1)\Delta}^{n\Delta} \Sigma_s^{1/2} dW_s \int_{(n-1)\Delta}^{n\Delta} dW_s^T \Sigma_s^{1/2} + \Delta \int_{(n-1)\Delta}^{n\Delta} \Sigma_s^{1/2} dW_s \mu^T \\ &\quad + \Delta \mu \int_{(n-1)\Delta}^{n\Delta} dW_s^T \Sigma_s^{1/2} - \Sigma_n, \end{aligned} \quad (4.45)$$

it follows that

$$\mathbf{Y}_n = \Delta\mu + u_{1,n} \quad (4.46)$$

$$\mathbf{Y}_n \mathbf{Y}_n^T = \Delta^2 \mu \mu^T + \Sigma_n + u_{2,n}. \quad (4.47)$$

To obtain a state space representation for (4.46) and (4.47) we need to derive the second order properties of  $u_n$ . Moreover, we express the increments of the integrated covariance  $\Sigma_n$  as a linear process with a noise sequence which is uncorrelated with  $u_n$ . The following proposition establishes the second order properties of  $u_n$ .

**Proposition 4.3.3.** *The sequence  $(u_n)_{n \in \mathbb{N}}$  is a (second order) stationary zero-mean martingale difference sequence (w.r.t. the filtration  $(\mathcal{G}_n)_{n \in \mathbb{N}} := (\mathcal{F}_{n\Delta})_{n \in \mathbb{N}}$ ) and thus in particular white noise. It holds that*

$$\text{var}(u_{1,n}) = E(\Sigma_n) = E(\Sigma_0)\Delta \quad (4.48)$$

$$\begin{aligned} \text{var}(\text{vec}(u_{2,n})) &= \Delta^2 (E(\Sigma_n) \otimes (\mu\mu^T) + (\mu\mu^T) \otimes E(\Sigma_n) + \mu^T \otimes E(\Sigma_n) \otimes \mu \\ &\quad + \mu \otimes E(\Sigma_n) \otimes \mu^T) + (I_{d^2} + \mathbf{P}) (E(\Sigma_n) \otimes E(\Sigma_n)) \\ &\quad + (\mathbf{Q} + \mathbf{PQ}) \text{var}(\text{vec}(\Sigma_n)) \end{aligned} \quad (4.49)$$

$$\begin{aligned} &= \Delta^3 (E(\Sigma_0) \otimes (\mu\mu^T) + (\mu\mu^T) \otimes E(\Sigma_0) + \mu^T \otimes E(\Sigma_0) \otimes \mu \\ &\quad + \mu \otimes E(\Sigma_0) \otimes \mu^T) + \Delta^2 (I_{d^2} + \mathbf{P}) (E(\Sigma_0) \otimes E(\Sigma_0)) \\ &\quad + (\mathbf{Q} + \mathbf{PQ}) (r^{++}(\Delta) + (r^{++}(\Delta))^T) \\ \text{cov}(u_{1,n}, \text{vec}(u_{2,n})) &= \Delta (E(\Sigma_n) \otimes \mu^T + \mu^T \otimes E(\Sigma_n)) \\ &= \Delta^2 (E(\Sigma_0) \otimes \mu^T + \mu^T \otimes E(\Sigma_0)) \end{aligned} \quad (4.50)$$

*Proof.* The proof of Proposition 4.3.3 is given in Pigorsch and Stelzer (2007).  $\square$

Note that so far the second order properties for  $u_n$  are derived for the general multivariate stochastic volatility. However, to express the increments of the integrated volatility as a linear process, the process of the volatility matrix needs to be specified in more detail. In the following we therefore turn our attention to the OU-type model and show that we can derive a state space representation for this specific model.

Defining

$$\eta_{1,n} := \int_{(n-1)\Delta}^{n\Delta} e^{A(n\Delta-s)} dL_s e^{A^T(n\Delta-s)} \quad (4.51)$$

$$\eta_{2,n} := \int_{(n-1)\Delta}^{n\Delta} dL_s = L_{n\Delta} - L_{(n-1)\Delta} \quad (4.52)$$

$$\eta_n := (\eta_{1,n}, \eta_{2,n}) \quad (4.53)$$

for all  $n \in \mathbb{N}$  it is easy to see that

$$\Sigma_{n\Delta} = e^{A\Delta} \Sigma_{(n-1)\Delta} e^{A^T \Delta} + \eta_{1,n} \quad (4.54)$$

$$L_{n\Delta} = L_{(n-1)\Delta} + \eta_{2,n}. \quad (4.55)$$

Before showing that this leads to a helpful state space representation we study the properties of the noise sequence  $(\eta_n)_{n \in \mathbb{N}}$ .



**Proposition 4.3.4.** *The sequence of random variables  $(\eta_n)_{n \in \mathbb{N}}$  is i.i.d. and uncorrelated with  $(u_n)_{n \in \mathbb{N}}$ . Moreover, it has finite second moments and*

$$\begin{aligned} E(\eta_{1,n}) &= -\mathbf{A}^{-1} \left( E(L_1) - e^{A\Delta} E(L_1) e^{A^T \Delta} \right) \\ &= E(\Sigma_0) - e^{A\Delta} E(\Sigma_0) e^{A^T \Delta} \end{aligned} \quad (4.56)$$

$$E(\eta_{2,n}) = \Delta E(L_1) = -\Delta \mathbf{A} E(\Sigma_0) \quad (4.57)$$

$$\begin{aligned} \text{var}(\text{vec}(\eta_{1,n})) &= -\mathcal{A}^{-1} \left( \text{var}(\text{vec}(L_1)) - e^{\mathcal{A}\Delta} \text{var}(\text{vec}(L_1)) e^{\mathcal{A}^T \Delta} \right) \\ &= \text{var}(\text{vec}(\Sigma_0)) - e^{\mathcal{A}\Delta} \text{var}(\text{vec}(\Sigma_0)) e^{\mathcal{A}^T \Delta} \end{aligned} \quad (4.58)$$

$$\text{var}(\text{vec}(\eta_{2,n})) = \Delta \text{var}(\text{vec}(L_1)) = -\Delta \mathcal{A} \text{var}(\text{vec}(\Sigma_0)) \quad (4.59)$$

$$\begin{aligned} \text{cov}(\text{vec}(\eta_{1,n}), \text{vec}(\eta_{2,n})) &= -\mathcal{A}^{-1} \left( \text{var}(\text{vec}(L_1)) - e^{\mathcal{A}\Delta} \text{var}(\text{vec}(L_1)) \right) \\ &= \mathcal{A}^{-1} \mathcal{A} (I_{d^2} - e^{\mathcal{A}\Delta}) \text{var}(\text{vec}(\Sigma_0)). \end{aligned} \quad (4.60)$$

*Proof.* The proof of Proposition 4.3.4 is given in Pigorsch and Stelzer (2007).  $\square$

Using the properties of the integrated variance (most importantly see Proposition 4.2.8) we have

$$\Sigma_n = \Sigma_{n\Delta}^+ - \Sigma_{(n-1)\Delta}^+ = \mathbf{A}^{-1} (\Sigma_{n\Delta} - \Sigma_{(n-1)\Delta} - L_{n\Delta} + L_{(n-1)\Delta})$$

for all  $n \in \mathbb{N}$  and recalling the definition of  $\eta_n$  we have

$$\mathbf{A} \Sigma_n = e^{A\Delta} \Sigma_{(n-1)\Delta} e^{A^T \Delta} - \Sigma_{(n-1)\Delta} + \eta_{1,n} - \eta_{2,n}.$$

Combining this with the representation (4.46) and (4.47) of the observable log-price process  $\mathbf{Y}_n$  and its “square”  $\mathbf{Y}_n \mathbf{Y}_n^T$ , and setting  $\alpha_{1,n} = \mathbf{A} \Sigma_n$  and  $\alpha_{2,n} = \Sigma_{n\Delta}$  results in the desired state space representation

$$\mathbf{Y}_n = \Delta \mu + u_{1,n} \quad (4.61)$$

$$\mathbf{Y}_n \mathbf{Y}_n^T = \Delta^2 \mu \mu^T + \mathbf{A}^{-1} \alpha_{1,n} + u_{2,n}. \quad (4.62)$$

where

$$\alpha_{1,n} = e^{A\Delta} \alpha_{2,n-1} e^{A^T \Delta} - \alpha_{2,n-1} + \eta_{1,n} - \eta_{2,n} \quad (4.63)$$

$$\alpha_{2,n} = e^{A\Delta} \alpha_{2,n-1} e^{A^T \Delta} + \eta_{1,n} \quad (4.64)$$

or in pure vector notation with  $\alpha_n := \begin{pmatrix} \text{vec}(\alpha_{1,n}) \\ \text{vec}(\alpha_{2,n}) \end{pmatrix}$ :

$$\begin{pmatrix} \mathbf{Y}_n \\ \text{vec}(\mathbf{Y}_n \mathbf{Y}_n^T) \end{pmatrix} = \begin{pmatrix} \Delta \mu \\ \Delta^2 (\mu \otimes \mu) \end{pmatrix} + \begin{pmatrix} 0_{M_{d,d^2}(\mathbb{R})} & 0_{M_{d,d^2}(\mathbb{R})} \\ \mathcal{A}^{-1} & 0_{M_{d^2,d^2}(\mathbb{R})} \end{pmatrix} \alpha_n + \begin{pmatrix} u_{1,n} \\ \text{vec}(u_{2,n}) \end{pmatrix} \quad (4.65)$$

where

$$\alpha_n = \begin{pmatrix} 0_{M_{d^2}(\mathbb{R})} & e^{A\Delta} \otimes e^{A\Delta} - I_{d^2} \\ 0_{M_{d^2}(\mathbb{R})} & e^{A\Delta} \otimes e^{A\Delta} \end{pmatrix} \alpha_{n-1} + \begin{pmatrix} \text{vec}(\eta_{1,n} - \eta_{2,n}) \\ \text{vec}(\eta_{1,n}) \end{pmatrix}. \quad (4.66)$$

Observe that  $0_{M_{d,d^2}(\mathbb{R})}$  denotes the zero matrix in  $M_{d,d^2}(\mathbb{R})$  etc.

### 4.3.3 Realized Quadratic Variation

In the following we derive a state space representation for the realized covariation. Recall from Section 2.1.2 that the realized covariation is defined as:

$$[\mathbf{Y}]_n^{(M)} := \sum_{j=1}^M \mathbf{Y}_{j,n} \mathbf{Y}_{j,n}^T. \quad (4.67)$$

In a first step we derive the second order properties of the realized variation.

**Theorem 4.3.5.** *Consider the general stochastic volatility model with  $\mu = \beta = 0$ . Then the sequence  $([\mathbf{Y}]_n^{(M)})_{n \in \mathbb{N}}$  is (second order) stationary and*

$$E([\mathbf{Y}]_n^{(M)} | \boldsymbol{\Sigma}_n) = \boldsymbol{\Sigma}_n \quad (4.68)$$

$$E([\mathbf{Y}]_n^{(M)}) = E(\boldsymbol{\Sigma}_n) = \Delta E(\Sigma_0) \quad (4.69)$$

$$\begin{aligned} \text{var}(\text{vec}([\mathbf{Y}]_n^{(M)})) &= \text{var}(\text{vec}(\boldsymbol{\Sigma}_n)) + M(\mathbf{Q} + \mathbf{Q}\mathbf{P}) \left( r^{++} \left( \frac{\Delta}{M} \right) + \left( r^{++} \left( \frac{\Delta}{M} \right) \right)^T \right) \\ &\quad + \frac{\Delta^2}{M} (I_{d^2} + \mathbf{P})(E(\Sigma_0) \otimes E(\Sigma_0)) \\ &= r^{++}(\Delta) + (r^{++}(\Delta))^T \\ &\quad + M(\mathbf{Q} + \mathbf{Q}\mathbf{P}) \left( r^{++} \left( \frac{\Delta}{M} \right) + \left( r^{++} \left( \frac{\Delta}{M} \right) \right)^T \right) \\ &\quad + \frac{\Delta^2}{M} (I_{d^2} + \mathbf{P})(E(\Sigma_0) \otimes E(\Sigma_0)) \\ \text{acov}_{[\mathbf{Y}]^{(M)}}(h) &= \text{acov}_{\boldsymbol{\Sigma}}(h) \\ &= r^{++}(h\Delta + \Delta) - 2r^{++}(h\Delta) + r^{++}(h\Delta - \Delta), \quad h \in \mathbb{N}. \end{aligned} \quad (4.70)$$

$$\text{acov}_{[\mathbf{Y}]^{(M)}}(h) = \text{acov}_{\boldsymbol{\Sigma}}(h) \quad (4.71)$$

*Proof.* The proof of Theorem 4.3.5 is given in Pigorsch and Stelzer (2007).  $\square$

Moreover, as the realized covariation is defined as the sum of the squared high-frequency returns, we can also generalize the state space representation of the last section by simple addition, whereby we have again to switch to the specific OU-type

stochastic volatility model with  $\beta = 0$ . We define for  $n \in \mathbb{N}$  and  $j = 1, 2, \dots, M$

$$u_n^{(M)} := \begin{pmatrix} u_{1,n}^{(M)} \\ u_{2,n}^{(M)} \end{pmatrix} \quad (4.72)$$

$$u_{1,j,n}^{(M)} := \int_{((n-1)+\frac{j-1}{M})\Delta}^{(n+\frac{j}{M})\Delta} \Sigma_s^{1/2} dW_s \quad (4.73)$$

$$\begin{aligned} u_{2,j,n}^{(M)} := & \int_{((n-1)+\frac{j-1}{M})\Delta}^{(n+\frac{j}{M})\Delta} \Sigma_s^{1/2} dW_s \int_{((n-1)+\frac{j-1}{M})\Delta}^{(n+\frac{j}{M})\Delta} dW_s^T \Sigma_s^{1/2} + \Delta \int_{((n-1)+\frac{j-1}{M})\Delta}^{(n+\frac{j}{M})\Delta} \Sigma_s^{1/2} dW_s \mu^T \\ & + \Delta \mu \int_{((n-1)+\frac{j-1}{M})\Delta}^{(n+\frac{j}{M})\Delta} dW_s^T \Sigma_s^{1/2} - \Sigma_{j,n} \end{aligned} \quad (4.74)$$

$$u_{1,n}^{(\mathcal{M})} := \sum_{j=1}^M u_{1,j,n}^{(M)} \quad (4.75)$$

$$u_{2,n}^{(\mathcal{M})} := \sum_{j=1}^M u_{2,j,n}^{(M)}. \quad (4.76)$$

Then the sequence  $u_{j,n}^{(M)} := (u_{1,j,n}^{(M)}, u_{2,j,n}^{(M)})$  is in law equal to the sequence  $u_n$  with  $\Delta$  replaced by  $\Delta/M$  and it is straightforward to see that

$$\mathbf{Y}_n = \Delta \mu + u_{1,n}^{(\mathcal{M})} \quad (4.77)$$

$$[\mathbf{Y}]_n^{(M)} = \frac{\Delta^2}{M} \mu \mu^T + \Sigma_n + u_{2,n}^{(\mathcal{M})} \quad (4.78)$$

and  $(u_n^{(\mathcal{M})})_{n \in \mathbb{N}} := (u_{1,n}^{(\mathcal{M})}, u_{2,n}^{(\mathcal{M})})_{n \in \mathbb{N}}$  is a (second order) stationary zero-mean martingale difference sequence with respect to the filtration  $\mathcal{G}_n := \mathcal{F}_{n\Delta}$ ,  $n \in \mathbb{N}$ . Moreover,

$$\text{var}(u_{1,n}^{(\mathcal{M})}) = E(\Sigma_n) = \Delta E(\Sigma_0) \quad (4.79)$$

$$\begin{aligned} \text{var}(\text{vec}(u_{2,n}^{(\mathcal{M})})) &= M \text{var}(\text{vec}(u_{2,1,n})) = \frac{\Delta^3}{M^2} (E(\Sigma_0) \otimes (\mu \mu^T) + (\mu \mu^T) \otimes E(\Sigma_0) \\ &\quad + \mu^T \otimes E(\Sigma_0) \otimes \mu + \mu \otimes E(\Sigma_0) \otimes \mu^T) \\ &\quad + \frac{\Delta^2}{M} (I_{d^2} + \mathbf{P}) (E(\Sigma_0) \otimes E(\Sigma_0)) \\ &\quad + M(\mathbf{Q} + \mathbf{PQ}) (r^{++}(\Delta/M) + (r^{++}(\Delta/M))^T) \end{aligned} \quad (4.80)$$

$$\begin{aligned} \text{cov}(u_{1,n}^{(\mathcal{M})}, \text{vec}(u_{2,n}^{(\mathcal{M})})) &= \frac{\Delta}{M} (E(\Sigma_n) \otimes \mu^T + \mu^T \otimes E(\Sigma_n)) \\ &= \frac{\Delta^2}{M} (E(\Sigma_0) \otimes \mu^T + \mu^T \otimes E(\Sigma_0)). \end{aligned} \quad (4.81)$$

In the OU-type stochastic volatility model we can use equations (4.63) and (4.64) to obtain a recursion for  $(\mathbf{A}\Sigma_n, \Sigma_{n\Delta})$ . Hence, we obtain the following state space

representation

$$\mathbf{Y}_n = \Delta\mu + u_{1,n}^{(\mathcal{M})} \quad (4.82)$$

$$[\mathbf{Y}]_n^{(\mathcal{M})} = \frac{\Delta^2}{M} \mu \mu^T + \alpha_{1,n} + u_{2,n}^{(\mathcal{M})} \quad (4.83)$$

where

$$\alpha_{1,n} = e^{A\Delta} \alpha_{2,n-1} e^{A^T \Delta} - \alpha_{2,n-1} + \eta_{1,n} - \eta_{2,n} \quad (4.84)$$

$$\alpha_{2,n} = e^{A\Delta} \alpha_{2,n-1} e^{A^T \Delta} + \eta_{1,n}. \quad (4.85)$$

Alternatively in pure vector notation with  $\alpha_n := \begin{pmatrix} \text{vec}(\alpha_{1,n}) \\ \text{vec}(\alpha_{2,n}) \end{pmatrix}$ :

$$\begin{pmatrix} \mathbf{Y}_n \\ \text{vec}([\mathbf{Y}]_n^{(\mathcal{M})}) \end{pmatrix} = \begin{pmatrix} \Delta\mu \\ \frac{\Delta^2}{M}(\mu \otimes \mu) \end{pmatrix} + \begin{pmatrix} 0_{M_{d,d^2}(\mathbb{R})} & 0_{M_{d,d^2}(\mathbb{R})} \\ \mathcal{A}^{-1} & 0_{M_{d^2,d^2}(\mathbb{R})} \end{pmatrix} \alpha_n + \begin{pmatrix} u_{1,n}^{(\mathcal{M})} \\ \text{vec}(u_{2,n}^{(\mathcal{M})}) \end{pmatrix} \quad (4.86)$$

where

$$\alpha_n = \begin{pmatrix} 0_{M_{d^2}(\mathbb{R})} & e^{A\Delta} \otimes e^{A\Delta} - I_{d^2} \\ 0_{M_{d^2}(\mathbb{R})} & e^{A\Delta} \otimes e^{A\Delta} \end{pmatrix} \alpha_{n-1} + \begin{pmatrix} \text{vec}(\eta_{1,n} - \eta_{2,n}) \\ \text{vec}(\eta_{1,n}) \end{pmatrix}. \quad (4.87)$$

Furthermore, recall that the sequence  $(\eta_n)_{n \in \mathbb{N}} := (\eta_{1,n}, \eta_{2,n})_{n \in \mathbb{N}}$  is i.i.d. and note that it is uncorrelated with  $(u_n^{(\mathcal{M})})_{n \in \mathbb{N}}$ . The further second order properties are given by equations (4.56) to (4.60).

## 4.4 Estimation Methods and Finite Sample Properties

Having derived the positive semidefinite OU–type stochastic volatility model, we now turn to its estimation. In particular, we provide a discussion of the different estimation methods along with a small Monte–Carlo analysis designed to assess their finite–sample properties.

### 4.4.1 Estimation Methods

As noted earlier, the estimation of continuous–time stochastic volatility models is complicated by the unavailability of the likelihood function. However, based on the theoretical results derived above, we can estimate multivariate OU–type stochastic volatility models either by using the second order dependence structure of the squared returns or by exploiting the different state space representations.

### Estimation via the second order dependence structure

Given the closed-form expression for the autocovariance function of the squared returns, see theorem 4.3.2, model estimates can be obtained by minimizing the distance between the empirical autocovariance function and the model-implied one, such that this estimation approach can be interpreted as a nonlinear and overidentified Yule–Walker estimator. In particular, our objective function is given by

$$\begin{aligned} SSR(i, p) = & \left\| \widehat{\text{var}} \left( \text{vec} \left( \mathbf{Y}(i) \mathbf{Y}(i)^T \right) \right) - \text{var} \left( \text{vec} \left( \mathbf{Y}(i) \mathbf{Y}(i)^T \right) \right) \right\| \\ & + \sum_{k=1}^{p/i} \left\| \widehat{\text{acov}}_{\mathbf{Y}(i) \mathbf{Y}(i)^T}(k) - \text{acov}_{\mathbf{Y}(i) \mathbf{Y}(i)^T}(k) \right\|, \end{aligned} \quad (4.88)$$

with  $i (= \Delta/M)$  denoting the sampling frequency, such as 15 minutes, and  $p$  the number of lags over time period  $\Delta$ , e.g. a day. Moreover, although several matrix norms can be considered we use the Frobenius norm, i.e.

$$\|A\| := \sum_{i=1}^d \sum_{j=1}^d a_{ij}^2 = \text{tr}(AA^T).$$

As noted in Barndorff-Nielsen and Shephard (2001b, Chapter 5.3), who apply this procedure for the estimation of a univariate OU-type stochastic volatility model, the estimator is independent of the assumption of a particular OU-type process. Instead, as can also be seen from the formulas of the autocovariance function 4.3.2, the estimator generally identifies the mean and the variance of the BDLP, as well as the matrix  $A$ . So, instead of assuming a specific parametric model for the BDLP, we optimize over the first two moments of  $L_1$ .<sup>4</sup>

### Estimation via the state space representations

Based on the state space representation for the squared high-frequency returns and for the realized covariation (see section 4.3.2, respectively) the Kalman filter can be used to obtain the quasi likelihood function of the model. Note that since we have not established the distributions of the volatility dynamics, i.e. we have not specified the distributions of  $u_n$  and  $\eta_n$ , the Kalman filter assumes Gaussianity and hence the resulting estimates are quasi maximum likelihood estimates. Both of these estimation approaches have also been considered in Barndorff-Nielsen and Shephard (2001b, 2002a) for the univariate OU-type stochastic volatility model, for which they also show, that the Kalman filter is suboptimal but provides consistent and asymptotically normal estimators. From the state space representations it also becomes clear, that the multivariate model can again be estimated without specifying a particular parametric BDLP. Instead we estimate the mean and the variance of the BDLP as well as the matrix  $A$ . For financial applications the

---

<sup>4</sup>Oftentimes the parameters of a specific BDLP can be identified solely by the mean and the variance of the BDLP. In this case, the autocovariance fit also identifies these parameters.

estimation via the Kalman filter might be preferable over the autocovariance fit, as it allows to compute smoothed and predicted series of the integrated covariance in a straightforward manner.

#### 4.4.2 Monte–Carlo Analysis

So far—at least to the best of our knowledge—a direct comparison of the three different estimation procedures in the univariate setting has not been pursued in the literature yet. In particular, we are not aware of any study presenting parameter estimates for the same dataset based on the different estimation methods, nor of any simulation study assessing the finite sample properties of the respective methods. This is the aim of this section.

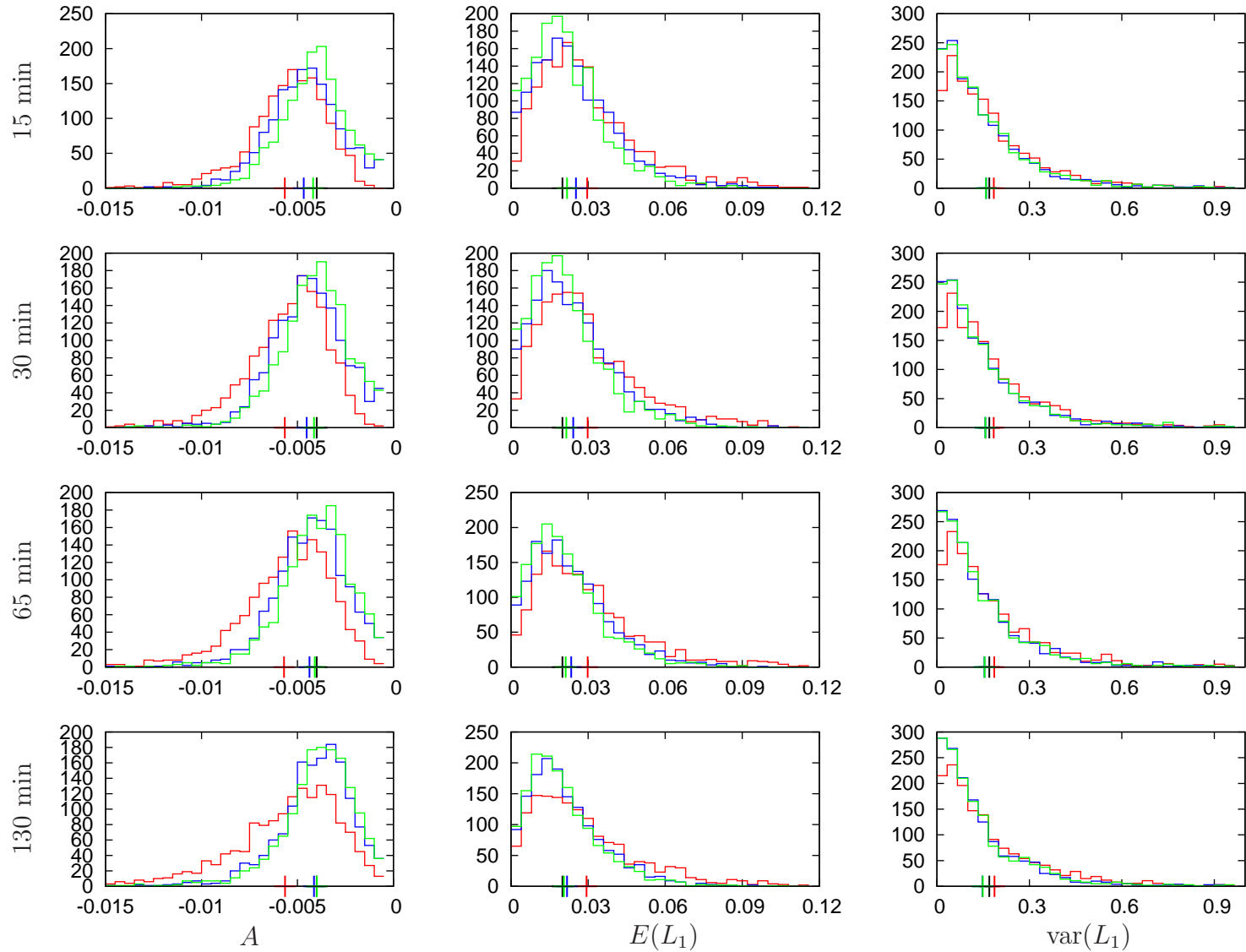
To this end, we generate a series of five minute returns over a period of 101250 days, from which we keep the last 1250 days (corresponding roughly to our sample size), from a univariate non–Gaussian OU–type process with model parameters set to empirically plausible values, i.e. they are based on our univariate estimation results for MSFT as reported in Section 4.2. For the simulated series we compute the 15, 30, 65 and 130 minute returns, the corresponding realized variation as well as the autocovariance function. The model is then estimated for these frequencies using the different estimation methods, whereby the autocovariance fit is based on  $p = 25$  daily lags. Note that the consideration of the different sampling frequencies allows us to assess the relevance of using more information. More specifically, for the estimation via the state space representation of the squared high–frequency returns as well as for the autocovariance fit an increasing sampling frequency leads to an increase in the number of observations used in the estimation. We therefore expect an improvement in the finite sample performance of this estimator at higher frequencies. In contrast, for the estimation based on the realized covariation the number of observations remains the same across different frequencies. However, the number of terms involved in the construction of the realized–covariation measure is increasing and so the resulting series should be more informative. We repeat our simulation procedure 1000 times providing us with a series of parameter estimates for  $A (= \frac{1}{2}\lambda)$ ,  $E(L_1)$  and  $\text{var}(L_1)$ . Given these estimates we can also compute for each simulation run some further statistics for which we have derived closed–form expression in the multivariate setting. In particular, we consider the variance of the daily returns and of the squared daily returns, as well as the autocorrelation coefficient of the squared returns at the 10th daily lag ( $\text{acorr}(y_t^2)_{10}$ ).

Table 4.1 presents some characteristics of the so obtained distribution of the respective statistics, which in turn are depicted in Figure 4.4 and 4.5. The results reveal, that the autocovariance–based estimation method tends to underestimate the parameter  $A$ , and to overestimate  $E(L_1)$  and  $\text{var}(y_t)$  irrespective of the sampling frequency used. Moreover, as expected the parameter estimates become slightly less precise with decreasing sampling frequency, i.e. with a decreasing number of observations used in the estimation. The two estimation methods based on the Kalman filter yield very accurate estimates, and only tend to underestimate the

**Table 4.1:** Monte–Carlo Results

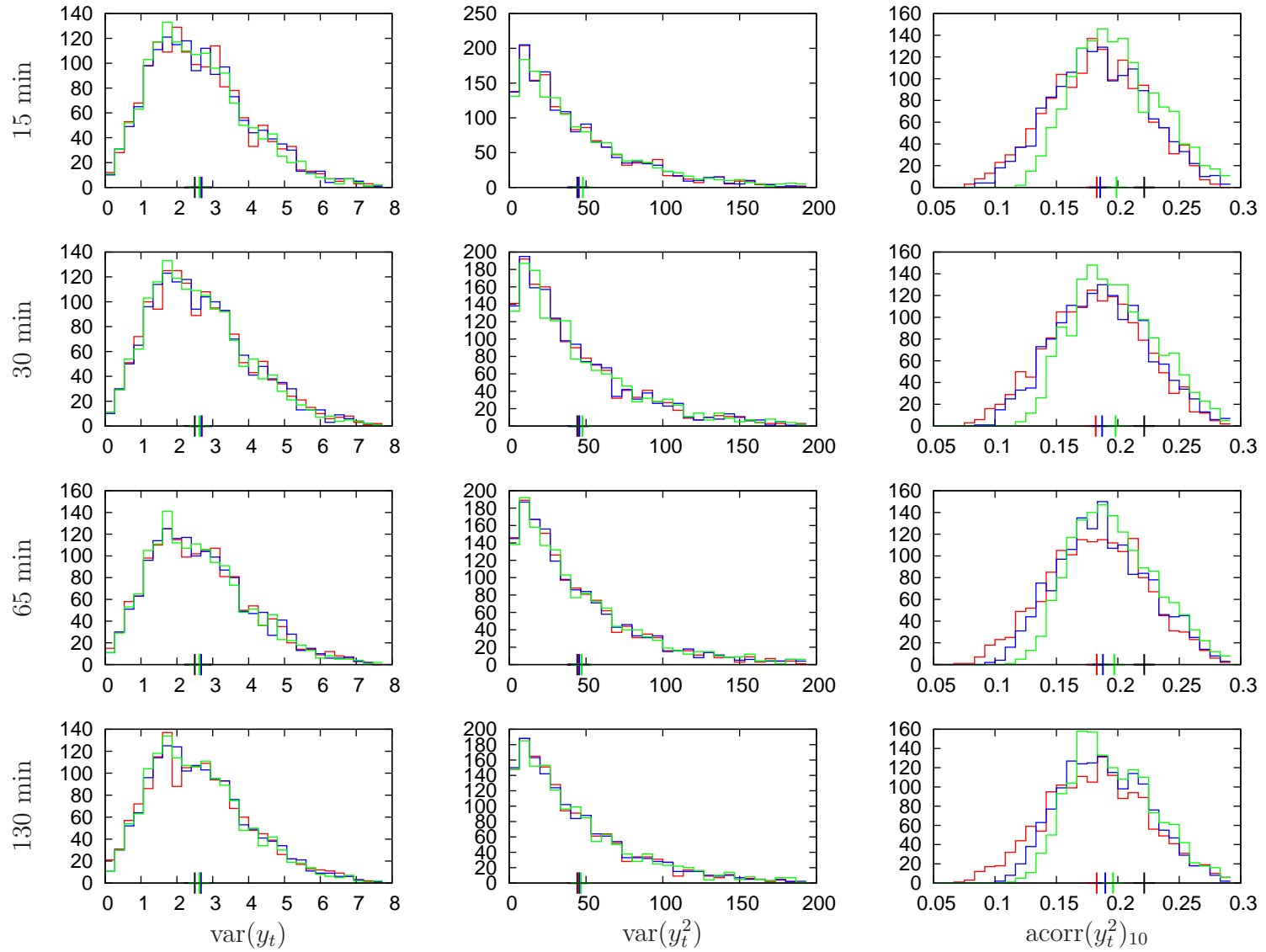
	autovoc				kalmanhf				kalmanrc			
	15min	30min	65min	130min	15min	30min	65min	130min	15min	30min	65min	130min
$A = -0.004$												
mean	-0.0057	-0.0057	-0.0057	-0.0057	-0.0047	-0.0045	-0.0044	-0.0041	-0.0042	-0.0041	-0.0041	-0.0040
std	0.0029	0.0029	0.0031	0.0034	0.0022	0.0022	0.0020	0.0020	0.0020	0.0020	0.0020	0.0021
min	-0.0638	-0.0636	-0.0616	-0.0540	-0.0403	-0.0411	-0.0250	-0.0179	-0.0306	-0.0382	-0.0258	-0.0201
max	-0.0010	-0.0007	-0.0000	-0.0000	-0.0004	-0.0004	-0.0004	-0.0004	-0.0003	-0.0004	-0.0003	-0.0003
$Q_{0.05}$	-0.0095	-0.0095	-0.0102	-0.0110	-0.0077	-0.0075	-0.0077	-0.0077	-0.0071	-0.0072	-0.0073	-0.0075
$Q_{0.50}$	-0.0053	-0.0052	-0.0053	-0.0050	-0.0046	-0.0044	-0.0042	-0.0039	-0.0041	-0.0040	-0.0039	-0.0037
$Q_{0.95}$	-0.0027	-0.0026	-0.0023	-0.0018	-0.0015	-0.0015	-0.0014	-0.0014	-0.0012	-0.0012	-0.0014	-0.0013
$E(L_1) = 0.02$												
mean	0.0297	0.0298	0.0298	0.0294	0.0253	0.0242	0.0234	0.0218	0.0218	0.0215	0.0213	0.0206
std	0.0196	0.0198	0.0210	0.0228	0.0170	0.0161	0.0158	0.0149	0.0144	0.0144	0.0144	0.0144
min	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0001	0.0000	0.0000	0.0001	0.0001
max	0.1301	0.1374	0.1454	0.1710	0.1237	0.1027	0.1222	0.1010	0.0900	0.0925	0.0912	0.1047
$Q_{0.05}$	0.0058	0.0059	0.0053	0.0044	0.0034	0.0034	0.0034	0.0033	0.0030	0.0032	0.0031	0.0033
$Q_{0.50}$	0.0255	0.0254	0.0255	0.0238	0.0220	0.0214	0.0202	0.0187	0.0194	0.0189	0.0186	0.0177
$Q_{0.95}$	0.0664	0.0677	0.0706	0.0727	0.0589	0.0563	0.0545	0.0503	0.0497	0.0501	0.0496	0.0480
$\text{var}(L_1) = 0.17$												
mean	0.1845	0.1844	0.1862	0.1869	0.1599	0.1572	0.1544	0.1475	0.1595	0.1571	0.1548	0.1488
std	0.1837	0.1843	0.1909	0.2093	0.1698	0.1668	0.1721	0.1645	0.1676	0.1670	0.1694	0.1672
min	0.0001	0.0000	0.0001	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
max	1.9006	1.6335	1.3110	2.0530	1.9998	1.6221	1.8559	1.2558	1.9753	1.5002	1.5041	1.2799
$Q_{0.05}$	0.0166	0.0164	0.0151	0.0130	0.0104	0.0102	0.0109	0.0096	0.0107	0.0103	0.0106	0.0097
$Q_{0.50}$	0.1355	0.1316	0.1285	0.1215	0.1124	0.1065	0.1026	0.0975	0.1137	0.1076	0.1036	0.0978
$Q_{0.95}$	0.5302	0.5320	0.5550	0.5520	0.4803	0.4598	0.4506	0.4359	0.4638	0.4619	0.4657	0.4340

Notes: The entries report the summary statistics of the simulated distributions of the parameter estimates based on the autocovariance–fit (autocov), on the Kalman filter for the squared high–frequency returns (kalmanhf), and on the Kalman filter for the realized covariation (kalmanrc) using different sampling frequencies. The first row in each panel reports the true parameter value.



**Figure 4.4:** Simulated distributions of the parameters estimates. The figure shows the simulated distributions of the parameter estimates based on the autocovariance-fit (red), the Kalman filter for the squared high-frequency returns (blue), and on the Kalman filter for the realized variation (green) using different frequencies (decreasing from upper to lower panel). The true parameter values are marked as black crosses, whereas the other crosses represent the mean of the parameter estimates colored according to the respective estimation method.





**Figure 4.5:** Simulated distributions of implied daily return characteristics. The figure shows the simulated distribution of the  $\text{var}(y_t)$ ,  $\text{var}(y_t^2)$  and the autocorrelation coefficient of the squared returns at the 10th lag ( $\text{acorr}(y_t^2)_{10}$ ) based on the estimation results of the autocovariance-fit (red), the Kalman filter for high-frequency returns (blue), and on the Kalman filter for the realized variation (green) using different sampling frequencies. The true parameter values are marked as black crosses, whereas the other crosses represent the mean of the parameter estimates colored according to the respective estimation method.

$\text{acorr}(y_t)_1 0$ . Moreover, for a decreasing sampling frequency the  $\text{var}(L_1)$  tends to be underestimated. Furthermore, for the realized–variation–based estimation the incorporation of more information by aggregating over more finely sampled returns does not seem to be essential for more precise parameter estimates. In particular, sampling at lower frequencies already provides very accurate results. Note also that the Kalman filter for the state space representation of the squared high–frequency returns seems to exhibit a slight tendency towards an underestimation of  $A$  and an overestimation of  $E(L_1)$  as the sampling frequency increases.

Overall the estimation via the state space representations yields more precise estimates than the autocovariance fit, whereby the estimates based on the realized variation are most stable across the different frequencies.

## 4.5 Empirical Application

In the following we estimate the positive semidefinite OU–type stochastic volatility model as given in equations (4.1) and (4.15), whereby we follow Barndorff-Nielsen and Shephard (2001b) by assuming that  $\mu = \beta = 0$ , i.e. the means of the returns are set to zero.<sup>5</sup> Moreover, we assume that the non–diagonal elements of the matrix subordinator are zero, such that we obtain in fact a vector subordinator. The elements of this subordinator are allowed to correlate positively. Hence, the correlation between the variances of the different assets is induced by both the correlation of the diagonal subordinators as well as the entries of  $A$ .

In order to assess the sensitivity of our model with respect to the sampling frequency, we estimate the model for different sampling frequencies. In particular, we consider 15 minutes, 30 minutes, 65 minutes and 130 minutes returns computed according to the description in Section 2.2.2.<sup>6</sup> Moreover, we assess the model’s performance for the raw (based only on the used trades as described in Section 2.2.2) as well as the adjusted dataset, whereby the adjustment is recomputed for each individual sampling frequency (i.e. in equation 2.20  $\Delta/M = 15, 30, 65$  and 130 minutes). This allows us to analyze the ability of our model to reproduce the seasonalities in the autocorrelation function induced by the intraday volatility pattern. So, overall, we consider a total of eight different datasets for which we estimate the model using the three different estimation methods discussed in the previous section.

Before presenting the estimation results of the multivariate Ornstein–Uhlenbeck–type stochastic volatility model, we first consider its univariate counterpart. This allows us a more detailed analysis of the multivariate model. We therefore start

---

<sup>5</sup>For a discussion on the estimation of these parameters either in a two–step approach or jointly using an extended state space representation of the model see Section 5 of Barndorff-Nielsen and Shephard (2001b). We expect that the proposed methods can also be applied in our multivariate case.

<sup>6</sup>Note, that we only consider frequency that are integer divisors of the total number of minutes of each trading day (i.e. 390 minutes). We therefore use 65 and 130 minutes rather than the oftentimes used 60 and 120 minutes frequency.

with a discussion of the univariate estimation results for Microsoft Corp. and Intel Corp.—the two assets for which we subsequently estimate the bivariate model.

Tables 4.2 and 4.3 present the univariate estimation results for Microsoft Corp. and Intel Corp., respectively, based on different sampling frequencies of the raw as well as the adjusted high-frequency data. The first five rows of each of the two panels for the raw and the adjusted data give the estimation results based on fitting the autocovariance function according to 4.88. The remaining entries correspond to the estimation via the Kalman filter using either the state space representation for the high-frequency returns or for the realized quadratic variation. Note, that based on the construction of realized quadratic variation the corresponding estimation results for the raw and the adjusted data are redundant and so we only report them in the raw data panel. The results show that irrespective of the estimation method a lower sampling frequency generally implies more persistence in the volatility process, i.e. smaller (absolute) values of  $A$ . At the same time, however, the mean (with the exception of the estimates based on the Kalman filter for realized variation) and the variance of the Background-driving Lévy process are also reduced. Hence, at lower frequencies the volatility is driven by on average smaller and less volatile jumps which are more persistent than at higher frequencies at which the jumps are larger and more volatile. Note, that the autocovariance-based estimates of  $A$  and  $\text{var}(L_1)$  are (in absolute terms) larger than the corresponding estimates of the Kalman filters, which is in accordance to our simulation results indicating that the parameter estimates based on the state space representations might be closer to the "true" ones. When evaluating the model's adequacy, we find that the models estimated via the state space representation can reproduce the empirical variance (measured over the whole sample period) of the returns, whereas the variance of the squared returns can hardly be captured by any of the estimation methods. In view of our simulation results, this suggests that our model is unable to reproduce adequately the fourth moments of the returns.

Figures 4.6 and 4.7 depict the estimated model-implied autocorrelations of the squared daily returns along with the empirical one (given by the straight line) for both, the raw (left panel) and the adjusted dataset (right panel) for Microsoft Corp. and Intel Corp. respectively. Note, that although the model is estimated for different sampling frequencies, we present the model-implied daily autocovariances (up to the 25th lag), in order to facilitate the comparison across the different estimation methods and sampling frequencies. In particular, the red lines refer to the daily autocorrelation functions implied by minimizing the distance between the empirical and the model-implied autocovariances at different intradaily frequencies. The blue and green lines depict the autocorrelation functions based on the estimation using the state space representation for the high-frequency returns and the realized quadratic variation, respectively. As can be seen from the figures, across the different estimation methods using a lower sampling frequency yields less persistence in the autocorrelation function than using a higher sampling frequency. At a first glance this might seem to be at odds with the findings in Tables 4.2 and 4.3, where we have observed a decreasing (absolute) value of  $A$  (implying more persistence) as

**Table 4.2:** Univariate Estimation Results for MSFT

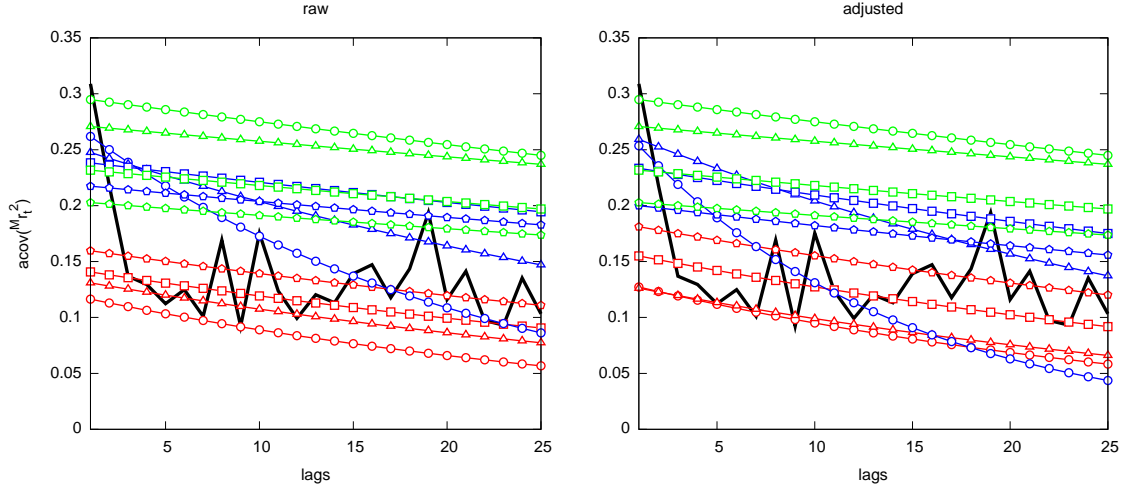
	$\Delta$	$A$	$E(L_1)$	$\text{var}(L_1)$	$\text{var}(r_t)$	$\text{var}(r_t^2)$
raw data					2.9957	48.5693
autocov	15min	-0.0150	0.1497	0.5583	5.0030	77.7664
	30min	-0.0110	0.0888	0.3204	4.0527	54.6189
	65min	-0.0092	0.0676	0.2499	3.6931	47.6199
	130min	-0.0076	0.0483	0.1928	3.1793	39.1680
kalmanhf	15min	-0.0232	0.1349	2.2730	2.9136	89.4921
	30min	-0.0108	0.0598	0.6768	2.7576	61.6699
	65min	-0.0042	0.0232	0.2174	2.7395	53.4075
	130min	-0.0036	0.0195	0.1336	2.6952	42.1717
kalmanrc	15min	-0.0039	0.0231	0.7424	2.9933	161.5467
	30min	-0.0028	0.0166	0.2933	2.9913	96.8841
	65min	-0.0034	0.0201	0.1854	2.9748	58.6616
	130min	-0.0032	0.0190	0.1179	2.9565	44.9131
adjusted data					2.9957	48.5693
autocov	15min	-0.0162	0.1567	0.6513	4.8226	76.2643
	30min	-0.0135	0.1197	0.4483	4.4387	64.1199
	65min	-0.0109	0.0768	0.3243	3.5216	46.9490
	130min	-0.0086	0.0485	0.2254	2.8385	35.7756
kalmanhf	15min	-0.0366	0.2141	3.3844	2.9216	84.6773
	30min	-0.0133	0.0734	1.0368	2.7676	73.4785
	65min	-0.0060	0.0329	0.2908	2.7491	51.3813
	130min	-0.0052	0.0280	0.1548	2.7010	36.8696

Notes: The estimation results are presented for the raw (upper panel) and the adjusted dataset (lower panel) and are based on different estimation methods: *autocov* refers to the estimation by fitting the model-implied autocovariance function (of up to the 25th lag) to the empirical one (see equation (4.88)); *kalmanhf* and *kalmanrc* refers to the estimation based on the state space representation for the high-frequency returns and for the realized variation, respectively.  $A$  is the mean-reversion parameter of the volatility process,  $E(L_1)$  and  $\text{var}(L_1)$  denote the mean and the variance, respectively, of the Background-driving Lévy process.  $\text{var}(r_t)$  and  $\text{var}(r_t^2)$  are the (full) sample variances of the daily returns and squared returns.

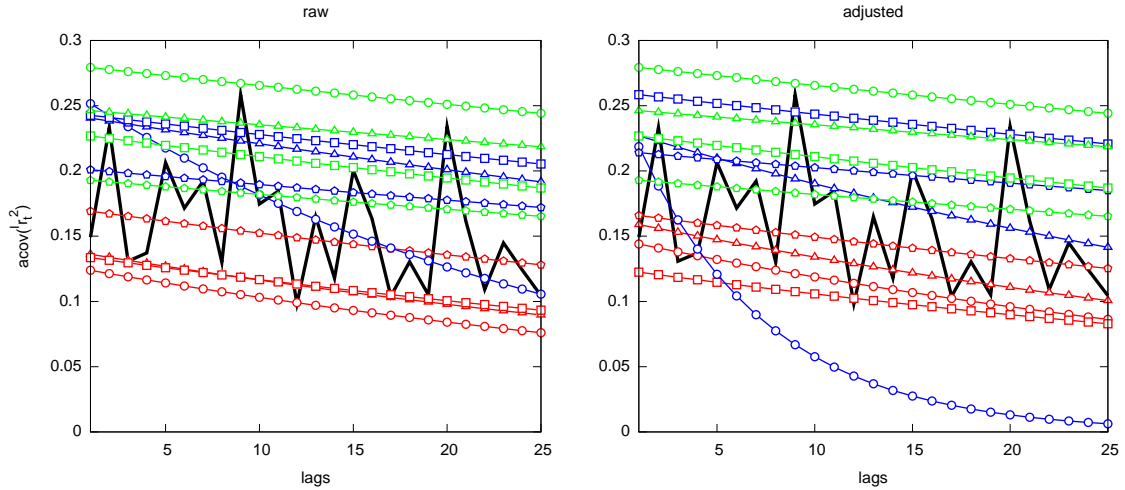
**Table 4.3:** Univariate Estimation Results for INTC

	$\Delta$	$A$	$E(L_1)$	$\text{var}(L_1)$	$\text{var}(r_t)$	$\text{var}(r_t^2)$
		raw data			5.6993	130.3840
autocov	15min	-0.0102	0.1721	1.1807	8.4630	230.7735
	30min	-0.0087	0.1261	0.8626	7.2494	179.0819
	65min	-0.0075	0.1115	0.7591	7.4713	187.5331
	130min	-0.0058	0.0685	0.5667	5.9060	142.7933
kalmanhf	15min	-0.0181	0.1973	4.9389	5.4513	261.7214
	30min	-0.0047	0.0499	0.9317	5.2725	202.9165
	65min	-0.0035	0.0373	0.7258	5.3532	213.3412
	130min	-0.0032	0.0351	0.3905	5.4244	149.1353
kalmanrc	15min	-0.0028	0.0319	1.2844	5.6930	407.8518
	30min	-0.0025	0.0283	0.6159	5.6926	250.3508
	65min	-0.0040	0.0455	0.7435	5.6502	202.0576
	130min	-0.0032	0.0360	0.3728	5.5852	148.9632
		adjusted data			5.6993	130.3840
autocov	15min	-0.0107	0.1635	1.3116	7.6575	208.7341
	30min	-0.0095	0.1251	1.0313	6.5766	167.3295
	65min	-0.0081	0.1360	0.9010	8.3650	222.5990
	130min	-0.0058	0.0731	0.6165	6.2614	157.2995
kalmanhf	15min	-0.0741	0.8087	16.1413	5.4562	215.1093
	30min	-0.0098	0.1040	1.6413	5.2844	180.0722
	65min	-0.0033	0.0354	0.8939	5.3728	261.0619
	130min	-0.0030	0.0330	0.4359	5.4473	166.9682

Notes: For a description of the entries reported in this table refer to Table 4.2.



**Figure 4.6:** Model-implied and empirical daily autocorrelation functions for MSFT. The figure shows the empirical (black straight line) daily autocorrelation function along with those implied by the parameter estimates based on the different estimation methods and sampling frequencies. The left panel presents the results for the raw data and the right panel those for the adjusted dataset. The autocorrelation functions based on the different sampling frequencies are characterized by different line styles, i.e.  $\circ$  refers to the 15 min sampling frequency,  $\triangle$  to the 30 min sampling frequency,  $\square$  to 65 min sampling frequency, and  $\diamond$  refers to the 130 min sampling frequency, whereas the different line colors represent a different estimation method. The red lines corresponds to the autocorrelation function implied by fitting the model-implied autocovariance function to the empirical one (see equation (4.88)), the blue lines corresponds to the estimation based on the state space representation for high-frequency returns, and the green lines corresponds to state space representation for the realized quadratic variation.



**Figure 4.7:** Model-implied and empirical daily autocorrelation functions for INTC. For the description of the figure refer to Figure 4.6.

the sampling frequency reduces. However, this can be explained by the expression of the autocorrelation function. In particular, for  $\mu = \beta = 0$  the autocorrelation function of our model is given in theorem 4.3.2 and does not only depend on  $A$  but also on the variance of the Background-driving Lévy process which has been found to be smaller for lower frequencies. The Figures also show, that estimating the model via the state space representation yields generally higher autocorrelations than the empirical one, whereby the estimation based on the realized quadratic variation is far-off for the 15 and 30 minutes sampling frequencies. For the remaining frequencies the two estimation methods perform pretty similar. Both approaches, however, are outperformed by the autocovariance estimation method, which is not surprising since it minimizes the difference of the model-implied and the empirical autocorrelation functions. The estimation results for the adjusted dataset exhibit similar characteristics as observed for the raw dataset, whereby the performance of the high-frequency state space based estimation is slightly improved (with the exception of the estimation using the 30 minutes frequency).

We now turn to the estimation of the bivariate positive semidefinite OU-type stochastic volatility model. As can be inferred from Tables 4.4 and 4.5 reporting the bivariate estimation results, the parameter estimates exhibit similar patterns as in the univariate case. In particular, the parameters estimates generally decrease with decreasing sampling frequencies, whereby the estimates based on the state space representation for the realized covariation are again most stable, and the autocovariance fits yield the largest variance and covariance estimates of the Background-driving Lévy processes. Interestingly, the sample variances and covariances of the daily returns are again most adequately reproduced by the estimation methods based on the Kalman filter.

**Table 4.4:** Bivariate Estimation Results for MSFT and INTC

	$\Delta$	$A_{1,1}$	$A_{2,1}$	$A_{2,2}$	$A_{2,2}$	$E(L_1)_{1,1}$	$E(L_1)_{2,2}$	$\text{var}(L_1)_{1,1}$	$\text{var}(L_1)_{2,1}$	$\text{var}(L_1)_{2,2}$
adjusted data										
autocov	15min	-0.0418	-0.0092	0.0641	-0.0058	0.1508	0.1289	2.3077	0.8506	0.4464
	30min	-0.0347	-0.0101	0.0559	-0.0009	0.1020	0.0779	1.3712	0.4777	0.2364
	65min	-0.0325	-0.0096	0.0533	0.0003	0.0716	0.0711	1.1874	0.4082	0.1880
	130min	-0.0301	-0.0118	0.0515	0.0039	0.0524	0.0464	0.9267	0.3176	0.1322
kalmanhf	15min	-0.0918	-0.0979	0.2081	0.0602	0.0122	0.1137	1.0260	0.4818	1.2163
	30min	-0.0289	-0.0309	0.0645	0.0215	0.0047	0.0234	0.3324	0.0814	0.1337
	65min	0.1641	0.1920	-0.3695	-0.1676	0.0000	0.0119	0.0498	0.0527	0.0559
	130min	0.1239	0.1458	-0.2845	-0.1269	0.0000	0.0102	0.0535	0.0354	0.0235
kalmanrc	15min	-0.0227	-0.0233	0.0423	0.0203	0.0103	0.0023	0.1640	0.0058	0.1484
	30min	-0.0218	-0.0219	0.0427	0.0183	0.0056	0.0090	0.1984	0.0000	0.1005
	65min	-0.0242	-0.0245	0.0497	0.0203	0.0000	0.0130	0.1557	0.0286	0.0482
	130min	-0.0223	-0.0231	0.0483	0.0190	0.0000	0.0111	0.1014	0.0193	0.0247
adjusted data										
autocov	15min	-0.0487	-0.0084	0.0727	-0.0081	0.0796	0.1409	2.8910	1.0450	0.5569
	30min	-0.0431	-0.0084	0.0662	-0.0055	0.0000	0.1113	1.9953	0.7124	0.3889
	65min	-0.0468	-0.0125	0.0764	0.0018	0.1091	0.0959	1.8047	0.5725	0.2346
	130min	-0.0355	-0.0127	0.0609	0.0051	0.0556	0.0509	1.1049	0.3623	0.1326
kalmanhf	15min	-3.0814	-3.6234	6.4958	3.0618	0.1282	0.0000	0.7329	0.4214	0.2423
	30min	0.5512	0.6492	-1.1240	-0.5574	0.0371	0.0000	0.1842	0.0967	0.0507
	65min	-0.0357	-0.0273	0.0670	0.0267	0.0575	0.0000	0.2301	0.0622	0.0576
	130min	-2.5726	-2.6987	5.6495	2.5680	0.0297	0.0000	0.0871	0.0410	0.0193

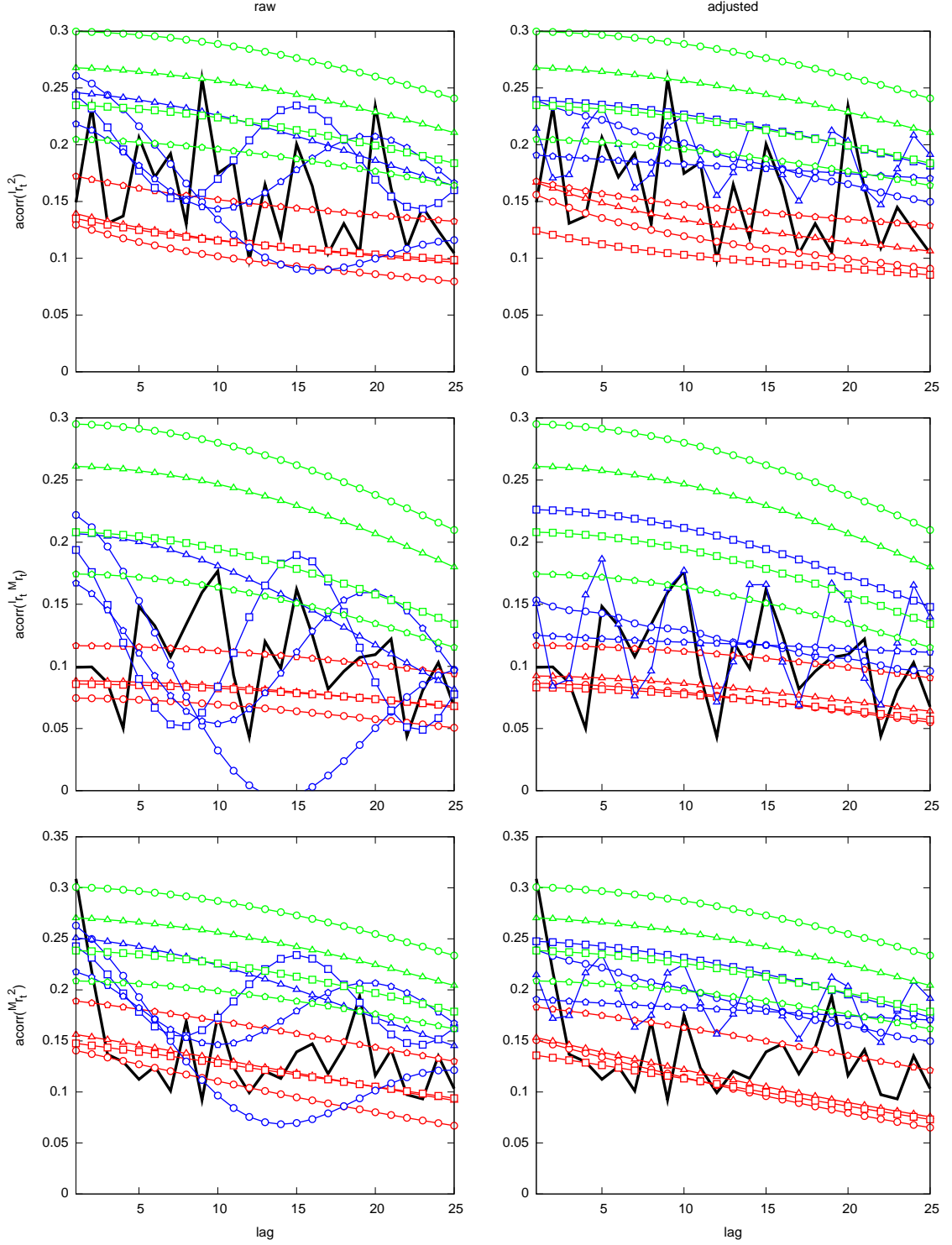
Notes: The estimation results are presented for the raw (upper panel) and the adjusted dataset (lower panel) and are based on different estimation methods: *autocov* refer to the estimation by fitting the model-implied autocovariance function (of up to the 25th lag) to the empirical one (see equation (4.88)); *kalmanrc* and *kalmanhf* refer to the estimation based on the state space representation for the high-frequency returns and for the realized covariation, respectively. Reported are the results for the entries of  $A$ , as well as of the mean vector and the covariance matrix of the Background-driving Lévy process.



**Table 4.5:** Bivariate estimation results of implied daily return characteristics (MSFT and INTC)

$\Delta$		$\langle^I r_t\rangle$	$\langle^I r_t,^M r_t\rangle$	$\langle^M r_t\rangle$	$\langle^I r_t^2\rangle$	$\langle^I r_t^2,^I r_t^M r_t\rangle$	$\langle^I r_t^2,^M r_t^2\rangle$	$\langle^I r_t^M r_t\rangle$	$\langle^M r_t^2,^I r_t^M r_t\rangle$	$\langle^M r_t^2\rangle$
raw data										
		5.6993	2.7390	2.9957	130.3840	68.8621	46.1189	55.6923	43.9153	48.5693
autocov	15min	8.3130	4.2474	4.3541	229.4144	103.8443	63.0617	81.2176	55.2869	66.3301
	30min	7.1490	3.5301	3.5459	177.7620	79.4417	47.6048	60.4924	40.5202	47.7711
	65min	7.4114	3.8396	3.6440	186.5218	87.2304	52.4303	64.6952	43.4620	47.9926
	130min	5.8185	2.8931	2.8034	141.9506	63.8522	39.2673	47.2088	31.5400	36.6306
kalmanhf	15min	5.4782	2.3887	2.9407	286.8854	122.8217	80.3240	90.7280	68.8467	90.6811
	30min	5.2632	2.3182	2.7865	212.5998	90.9424	58.2471	67.5391	49.5890	63.4889
	65min	5.2708	2.3405	2.7167	215.4310	95.7749	59.4672	68.3087	49.2744	56.8637
	130min	5.2825	2.2998	2.6826	164.6216	71.8231	42.7897	51.6716	36.4135	42.1818
kalmanrc	15min	5.4319	2.7953	3.1489	588.9409	312.0233	219.0245	228.3155	184.1286	205.2037
	30min	5.5893	2.7911	3.1045	318.9696	158.8653	106.5755	116.1370	90.5030	102.9625
	65min	5.7404	2.7881	3.0441	223.6810	106.3295	67.4487	77.1494	57.3190	65.4008
	130min	5.7383	2.6494	2.9239	171.2418	77.5129	46.4074	56.1662	40.0122	45.8656
adjusted data										
autocov	15min	7.2597	4.3172	4.2144	203.0017	99.2069	66.7592	78.7169	56.6038	65.8486
	30min	6.3516	4.1338	3.8360	165.1680	85.5240	60.0587	67.3353	49.5818	55.1687
	65min	8.3247	4.3853	3.9505	223.4970	107.8377	64.1211	77.7772	51.7927	53.0818
	130min	6.2097	3.1574	2.8644	157.4716	72.6837	44.0276	51.8453	34.1090	36.7130
kalmanhf	15min	5.4536	2.5771	3.0498	220.1967	104.2200	67.2668	77.2580	58.2839	68.8674
	30min	5.1923	2.5627	2.9845	198.0647	97.4740	63.9301	72.8592	55.9878	65.2553
	65min	5.8075	2.6695	2.7266	239.9187	120.5099	73.5427	82.2512	58.5539	58.1607
	130min	5.6639	2.5765	2.7077	150.7523	68.6008	39.0351	47.7327	32.7956	34.4539

Notes: The table presents the entries of the covariance matrix of the returns (first three columns) and of the variance of the squared returns (fourth and last column), of the crossproduct of the the returns and the squared returns of the two assets (column 7 and column 6, respectively), and of the variance of the crossproduct of the squared returns of one asset with the crossproduct of the returns of both assets (fifth column). The first row of each column reports the corresponding empirical values. For further descriptions of the entries reported in this table refer to Table 4.4.



**Figure 4.8:** Model-implied and empirical daily autocorrelation functions based on the bivariate estimation results for MSFT and INTC. The figure shows the model-implied and empirical (black line) autocorrelation functions of the squared returns of the two assets (upper and lower panel) as well as for their crossproduct (middle panel). For a further description of the figure refer to Figure 4.6.

Figure 4.8 shows the estimated as well as the empirical (in black) autocorrelation functions of the squared daily returns and the cross product of the daily returns of the two assets. Just as in the univariate case the Figure illustrates that a lower sampling frequency yields less persistence in the autocorrelation function across all estimation methods. Importantly, however, we can observe oscillations in the autocorrelation functions of the squared returns and the cross product implied by the state space estimation for the high-frequency returns. This might be caused by the ability of our multivariate model to reproduce the sinusoidal behavior in the intraday autocorrelation function induced by the U-shaped intraday volatility pattern (see e.g. Figure 2.6 of Section 2). Noteworthy, these oscillations disappear if the model is fitted to the intradaily-pattern adjusted dataset (with the exception of the autocorrelation function based on 30 minutes returns, which in fact did not exhibit such oscillations for the raw dataset, but is now quite accurate).

So, based on our simulation and estimation results, is there any estimation method that might be preferable in practice? The answer to this question depends on the application at hand. In particular, if the main focus is on an adequate reproduction of the autocorrelation function of the squared returns, then the autocovariance-based estimation method should be used. In contrast, if the main interest is on the integrated covariance matrix, as is the case in most financial risk management applications, then the estimation via the state space representations should be preferred, whereby our results indicate that the one for the realized covariation provides more robust results across different sampling frequencies. Moreover, in our empirical application these robust estimation results, i.e. the estimates based on the realized covariance, suggest that our multivariate OU-type stochastic volatility model can adequately reproduce the integrated covariance, but fails to account for the fourth moments in the return series.

## 4.6 Summary

Given the relevance of a joint modeling of the dynamics of multiple assets for portfolio and risk management decisions, we have extended the non-Gaussian Ornstein–Uhlenbeck-type stochastic volatility model Barndorff-Nielsen and Shephard (2001b) to the multivariate case. It turns out, that our model possesses many attractive features which are mainly a result of our stochastic volatility specification. In particular we have shown, that similar to Barndorff-Nielsen and Shephard (2001b) our general multivariate specification implies that returns are scaled mixtures of multivariate normals with the scaling given by the integrated covariance matrix, such that the observed fat-tailedness, as well as the volatility clustering of returns can be reproduced. Moreover, returns also aggregate to Gaussianity as the time interval over which returns are computed increases. In addition, specifying the stochastic volatility by Lévy-driven positive semidefinite Ornstein–Uhlenbeck-type processes provides a relatively flexible volatility dependency structure. In particular, we show, that the increments of the integrated covariance of a stochastic model

based on a single positive semidefinite Ornstein–Uhlenbeck–type process follows a VARMA(1,1) process. Furthermore, the first and second moments of the integrated volatility increments exhibit closed–form expressions. These results have several important implications. First, the implementation of financial decisions, such as the choice of e.g. a minimum–variance portfolio or other types of risk assessment, is facilitated. Second, based on the VARMA structure of the integrated covariance, we have derived a state space representation for the model–implied realized quadratic variation, which in turn can be used for the estimation of these models as well as for extracting and predicting the integrated covariance using the Kalman filter. Alternatively, a state space representation for the outer product of high–frequency returns can be used, which is based on the finding that those follow also a VARMA(1,1) process. Moreover, we derived closed–form expressions for the autocovariance function of the squared returns, which provides an additional approach for the estimation of our model. However, in practice the estimation based on the Kalman filter might be preferable as it provides quite accurate estimates of the integrated covariance matrix, which is of major importance for many financial applications.

## 5 Conclusion

In this thesis we analyze the usefulness of high-frequency financial data for the estimation and empirical assessment of different continuous-time stochastic volatility models. In doing so, we primarily exploit the realized variation measures and show that the information inherent in these measures does not only facilitate the statistical assessment of univariate as well as multivariate continuous-time stochastic volatility models, but also provides new insights into the empirical properties implied by the different models.

Motivated by the inability of the existing empirical studies based on daily data to allow for a clear distinction between pure diffusion multi-factor stochastic volatility models and lower-order models with jumps, we re-assess the adequacy of the affine and logarithmic jump-diffusion models within a unified framework. In particular, we employ the general scientific modeling method of Gallant and McCulloch (2005), which allows the model assessment of rather diverse structural models in terms of a highly accurate auxiliary model. To this end, we develop a discrete-time simultaneous equation model for the daily returns and realized variation measures, in which the dynamics of the total price variation is explicitly decomposed into its two components, i.e. the variation coming from the continuous-sample path evolution and the variation coming from the jumps. As a result we obtain a model that is highly informative on the most important empirical features that should be accounted for by the different continuous-time stochastic volatility models. Importantly, we show that the often observed leverage effect, or asymmetry in the relationship between lagged and contemporaneous returns and the volatility, primarily acts through the continuous volatility component, and that jumps are important and exhibit only weak own serial correlation.

Using the likelihood function of this auxiliary model, we estimate and assess the different continuous-time stochastic volatility models. Our estimation and simulation results reveal that the affine and logarithmic jump-diffusion models considered here still miss some important features of the data. In particular they are unable to reproduce the persistence and the leverage effects. More specifically, the one-factor models tend to underestimate the persistence whereas the two-factor model and the one-factor model with jumps tend to an overestimation but are able to reproduce a wider range of the persistence. Moreover, based on a conditional model assessment we find that the latter models imply more accurate tails of the return distributions than the one-factor models, suggesting that at least two volatility factors or one factor and a jump process are needed to better account for the volatility persistence and the fat-tailedness of the return distribution. Moreover, the inclusion of jumps into the price process leads to an improvement of the model's fit. Overall,

our results are novel to the literature in that sense that they provide more detailed information on the (in)ability of the models under consideration to capture specific empirical data characteristics. These findings could only be obtained through the use of high-frequency data.

Apart from the assessment of the individual price processes, this thesis also considers the joint modeling of multiple assets, which is important for portfolio and risk management decisions. In addition, based on the unavailability of a closed-form expression for the integrated covariance matrix of the existing multivariate continuous-time stochastic volatility models, we develop the multivariate positive semidefinite Ornstein–Uhlenbeck-type stochastic volatility model, for which the integrated covariance matrix is readily available and in fact exhibits a very simple structure. Moreover, the theoretical properties of this model allows us to derive three different estimation methods that are based on the use of high-frequency information, either using the high-frequency returns directly or the realized covariation measure. The Monte–Carlo study analyzing the small sample properties of the different estimation methods and our empirical application reveal that the most robust results are obtained by using the state space representation of the realized covariation measure. Moreover, the two estimations methods based on the state space representations either for the squared high-frequency returns or for the realized covariation measure might be preferable in practice, as the Kalman filter allows the computation of the smoothed and predicted series of the integrated covariance in a straightforward manner.

# Bibliography

- Aït-Sahalia, Y., Mykland, P. A., and Zhang, L. (2005), “How Often to Sample a Continuous-Time Process in the Presence of Market Microstructure Noise,” *Review of Financial Studies*, 18, 351–416.
- Andersen, T. G., Benzoni, L., and Lund, J. (2002), “An Empirical Investigation of Continuous-Time Equity Return Models,” *Journal of Finance*, 57, 1239–1284.
- Andersen, T. G. and Bollerslev, T. (1997), “Intraday Periodicity and Volatility Persistence in Financial Markets,” *Journal of Empirical Finance*, 4, 115–158.
- (1998), “Answering the Skeptics: Yes, Standard Volatility Models do Provide Accurate Forecasts,” *International Economic Review*, 39, 885–905.
- Andersen, T. G., Bollerslev, T., and Diebold, F. X. (2007), “Roughing it Up: Including Jump Components in the Measurement, Modeling and Forecasting of Return Volatility,” *Review of Economics and Statistics*, forthcoming.
- Andersen, T. G., Bollerslev, T., Diebold, F. X., and Ebens, H. (2001a), “The Distribution of Realized Stock Return Volatility,” *Journal of Financial Economics*, 61, 43–76.
- Andersen, T. G., Bollerslev, T., Diebold, F. X., and Labys, P. (2001b), “The Distribution of Realized Exchange Rate Volatility,” *Journal of the American Statistical Association*, 96, 42–55.
- (2003a), “Modeling and Forecasting Realized Volatility,” *Econometrica*, 71, 579–625.
- Andersen, T. G., Bollerslev, T., Diebold, F. X., and Vega, C. (2003b), “Micro Effects of Macro Announcements: Real-Time Price Discovery in Foreign Exchange,” *American Economic Review*, 93, 38–62.
- Andersen, T. G., Bollerslev, T., and Huang, X. (2006), “A Semiparametric Framework for Modeling and Forecasting Jumps and Volatility in Speculative Prices,” Working Paper, Duke University.
- Bachelier, L. (1900), “Théorie de la spéculation,” *Annales Scientifiques de l'École Normale Supérieure*, 3, 21–86, english translation in The random character of stock market prices, Ed. P. Cootner, pp. 17–78, Cambridge, MIT Press, 1964.

## Bibliography

- Bakshi, G., Cao, C., and Chen, Z. (1997), “Empirical Performance of Alternative Option Pricing Models,” *Journal of Finance*, 52, 2003–2049.
- Bandi, F. M. and Russell, J. R. (2005a), “Microstructure Noise, Realized Volatility, and Optimal Sampling,” Working Paper, The University of Chicago.
- (2005b), “Realized Covariation, Realized Beta, and Microstructure Noise,” Working Paper, The University of Chicago.
- Barndorff-Nielsen, O. E., Hansen, P. R., Lunde, A., and Shephard, N. (2006a), “Designing Realised Kernels to Measure the ex-post Variation of Equity Prices in the Presence of Noise,” Working Paper, University of Oxford.
- (2006b), “Subsampled Realised Kernels,” Working Paper, University of Oxford.
- Barndorff-Nielsen, O. E., Nicolato, E., and Shephard, N. (2002), “Some Recent Developments in Stochastic Volatility Modelling,” *Quantitative Finance*, 2, 11–23.
- Barndorff-Nielsen, O. E. and Pérez-Abreu, V. (2007), “Matrix Subordinators and Related Upsilon Transformations,” *Theory of Probability and its Applications*, forthcoming.
- Barndorff-Nielsen, O. E. and Shephard, N. (2001a), “Modelling by Lévy Processes for Financial Econometrics,” in *Lévy Processes – Theory and Applications*, eds. Barndorff-Nielsen, O. E., Mikosch, T., and Resnick, S. I., Basel: Birkhäuser, pp. 283–318.
- (2001b), “Non-Gaussian Ornstein–Uhlenbeck-based Models and Some of their Uses in Financial Economics,” *Journal of the Royal Statistical Society, Series B*, 63, 167–241.
- (2002a), “Econometric Analysis of Realised Volatility and its Use in Estimating Stochastic Volatility Models,” *Journal of the Royal Statistical Society, Series B*, 64, 253–280.
- (2002b), “Estimating Quadratic Variation Using Realized Variance,” *Journal of Applied Econometrics*, 17, 457–477.
- (2004a), “Econometric Analysis of Realized Covariation: High Frequency Based Covariance, Regression, and Correlation in Financial Economics,” *Econometrica*, 72, 885–925.
- (2004b), “Power and Bipower Variation with Stochastic Volatility and Jumps,” *Journal of Financial Econometrics*, 2, 1–37.



## Bibliography

- (2005), “How Accurate is the Asymptotic Approximation to the Distribution of Realised Variance?” in *Identification and Inference for Econometric Models. A Festschrift for Tom Rothenberg*, eds. Andrews, D., Powell, J., Ruud, P., and Stock, J., Cambridge: Cambridge University Press, pp. 306–331.
- (2006a), “Econometrics of Testing for Jumps in Financial Economics using Bipower Variation,” *Journal of Financial Econometrics*, 4, 1–30.
- (2006b), “Impact of Jumps on Returns and Realised Variances: Econometric Analysis of Time-Deformed Lévy Processes,” *Journal of Econometrics*, 131, 217–252.
- Barndorff-Nielsen, O. E. and Stelzer, R. (2006), “Roots of Positive-Definite Matrix Valued Processes and Ornstein-Uhlenbeck Processes,” Research Report 11, Thiele Centre, Department of Mathematical Sciences, University of Århus, Århus, Denmark.
- Bates, D. S. (1996a), “Jumps and Stochastic Volatility: Exchange Rate Processes Implicit in Deutsche Mark Options,” *Review of Financial Studies*, 9, 69–107.
- (1996b), “Testing Option Pricing Models,” in *Statistical Methods in Finance*, eds. Maddala, G. and Rao, C. R., Elsevier, vol. 14 of *Handbook of Statistics*, pp. 567–611.
- (2000), “Post-’87 Crash Fears in the S&P 500 Futures Option Market,” *Journal of Econometrics*, 94, 181–238.
- Bekaert, G. and Wu, G. (2000), “Asymmetric Volatility and Risk in Equity Markets,” *Review of Financial Studies*, 13, 1–42.
- Bibby, B. M. and Sørensen, M. (2001), “Simplified Estimating Functions for Diffusion Models with a High-dimensional Parameter,” *Scandinavian Journal of Statistics*, 28, 99–112.
- Black, F. and Scholes, M. (1973), “The Pricing of Options and Corporate Liabilities,” *Journal of Political Economy*, 81, 637–654.
- Bollen, B. and Inder, B. (2002), “Estimating Daily Volatility in Financial Markets Utilizing Intraday Data,” *Journal of Empirical Finance*, 9, 551–562.
- Bollerslev, T., Engle, R., and Nelson, D. (1994), “ARCH Models,” in *Handbook of Econometrics*, eds. Engle, R. F. and McFadden, D., Elsevier, chap. 49, pp. 2959–3038.
- Bollerslev, T. (1987), “A Conditionally Heteroskedastic Time Series Model for Speculative Prices and Rates of Return,” *Review of Economics and Statistics*, 69, 542–547.

## Bibliography

- Bollerslev, T., Cai, J., and Song, F. M. (2000), “Intraday Periodicity, Long-Memory Volatility, and Macroeconomic Announcement Effects in the U.S. Treasury Bond Market,” *Journal of Empirical Finance*, 7, 37–55.
- Bollerslev, T., Engle, R. F., and Wooldridge, J. (1988), “A Capital Asset Pricing Model with Time Varying Covariances,” *Journal of Political Economy*, 96, 116–131.
- Bollerslev, T., Gallant, R., Pigorsch, C., Pigorsch, U., and Tauchen, G. (2006a), “Statistical Assessment of Models for Very High Frequency Financial Price Dynamics,” Working Paper, Duke University.
- Bollerslev, T., Kretschmer, U., Pigorsch, C., and Tauchen, G. (2007), “A Discrete-Time Model for Daily S&P500 Returns and Realized Variations: Jumps and Leverage Effects,” Working Paper, Duke University.
- Bollerslev, T., Litvinova, J., and Tauchen, G. (2006b), “Leverage and Volatility Feedback Effects in High-Frequency Data,” *Journal of Financial Econometrics*, 4, 353–384.
- Bollerslev, T. and Zhou, H. (2002), “Estimating Stochastic Volatility Diffusions Using Conditional Moments of Integrated Volatility,” *Journal of Econometrics*, 109, 33–65.
- Brockwell, P. (2001), “Lévy-Driven CARMA Processes,” *Annals of the Institute of Statistical Mathematics*, 53, 113–124.
- Campbell, J. Y. and Hentschel, L. (1992), “No News is Good News: An Asymmetric Model of Changing Volatility in Stock Returns,” *Journal of Financial Economics*, 31, 281–331.
- Carr, P., Geman, H., Madan, D. B., and Yor, M. (2003), “Stochastic Volatility for Lévy Processes,” *Mathematical Finance*, 13, 345–382.
- Chernov, M., Gallant, A. R., Ghysels, E., and Tauchen, G. (2003), “Alternative Models for Stock Price Dynamics,” *Journal of Econometrics*, 116, 225–257.
- Chib, S., Nardari, F., and Shephard, N. (2006), “Analysis of High-Dimensional Multivariate Stochastic Volatility Models,” *Journal of Econometrics*, 134, 341–371.
- Comte, F. and Renault, E. (1998), “Long-Memory in Continuous-Time Stochastic Volatility Models,” *Mathematical Finance*, 8, 291–323.
- Corsi, F. (2004), “A Simple Long Memory Model of Realized Volatility,” Working Paper, University of Southern Switzerland.

## Bibliography

- Corsi, F., Pigorsch, U., Mitnik, S., and Pigorsch, C. (2007), “The Volatility of Realized Volatility,” *Econometric Reviews*, forthcoming.
- Dai, Q. and Singleton, K. J. (2000), “Specification Analysis of Affine Term Structure Models,” *Journal of Finance*, 55, 1943–1978.
- Davydov, Y. A. (1973), “Mixing Conditions for Markov Chains,” *Theory of Probability and its Applications*, 18, 312–328.
- Deo, R., Hurvich, C., and Lu, Y. (2006), “Forecasting Realized Volatility Using a Long-Memory Stochastic Volatility Model: Estimation, Prediction And Seasonal Adjustment,” *Journal of Econometrics*, 131, 29–58.
- Duffie, D., Pan, J., and Singleton, K. (2000), “Transform Analysis and Asset Pricing for Affine Jump-diffusions,” *Econometrica*, 68, 1343–1376.
- Duffie, D. and Singleton, K. J. (1993), “Simulated Moments Estimation of Markov Models of Asset Prices,” *Econometrica*, 61, 929–952.
- Durham, G. B. and Gallant, A. R. (2002), “Numerical Techniques for Maximum Likelihood Estimation of Continuous-Time Diffusion Processes,” *Journal of Business & Economic Statistics*, 20, 297–316.
- Engle, R. F. and Lee, G. (1999), “A Permanent and Transitory Component Model of Stock Return Volatility,” *Cointegration, Causality, and Forecasting: A Festschrift in Honor of Clive W. J. Granger*, 475–497.
- Engle, R. F. and Ng, V. K. (1993), “Measuring and Testing the Impact of News on Volatility,” *Journal of Finance*, 48, 1749–1778.
- Engle, R. F. and Russell, J. R. (2007), “Analysis of High-Frequency and Transaction Data,” in *Handbook of Financial Econometrics*, eds. Aït-Sahalia, Y. and Hansen, L. P., Elsevier, p. forthcoming.
- Eraker, B. (2001), “MCMC Analysis of Diffusion Models with Application to Finance,” *Journal of Business & Economic Statistics*, 19, 177–191.
- Eraker, B., Johannes, M., and Polson, N. (2003), “The Impact of Jumps in Volatility and Returns,” *Journal of Finance*, 58, 1269–1300.
- Forsberg, L. and Bollerslev, T. (2002), “Bridging the Gap Between the Distribution of Realized (ECU) Volatility and ARCH Modeling (of the Euro): The GARCH-NIG Model,” *Journal of Applied Econometrics*, 17, 535–548.
- French, K. R., Schwert, G. W., and Stambaugh, R. F. (1987), “Expected Stock Returns and Volatility,” *Journal of Financial Economics*, 19, 3–29.

## Bibliography

- Gallant, A. R. and McCulloch, R. E. (2005), “On the Determination of General Scientific Models with Application to Asset Pricing,” Working Paper, Duke University.
- Gallant, A. R. and Nychka, D. W. (1987), “Semi-Nonparametric Maximum Likelihood Estimation,” *Econometrica*, 55, 363–390.
- Gallant, A. R. and Tauchen, G. (1989), “Seminonparametric Estimation of Conditionally Constrained Heterogeneous Processes: Asset Pricing Applications,” *Econometrica*, 57, 1091–1120.
- (1996), “Which Moment to Match?” *Econometric Theory*, 12, 657–681.
- Ghysels, E., Santa-Clara, P., and Valkanov, R. (2005), “There is a Risk-Return Trade-Off After All,” *Journal of Financial Economics*, 76, 509–548.
- Gourieroux, C. (2006), “Continuous Time Wishart Process for Stochastic Risk,” *Econometric Reviews*, 25, 177–217.
- Gourieroux, C., Monfort, A., and Renault, E. (1993), “Indirect Inference,” *Journal of Applied Econometrics*, 8, S85–S118.
- Hansen, L. P. and Scheinkman, J. A. (1995), “Back to the Future: Generating Moment Implications for Continuous-Time Markov Processes,” *Econometrica*, 63, 767–804.
- Hansen, P. R., Large, J. H., and Lunde, A. (2007), “Moving Average Based Estimators of Integrated Variance,” *Econometric Reviews*, forthcoming.
- Hansen, P. R. and Lunde, A. (2006), “Consistent Ranking of Volatility Models,” *Journal of Econometrics*, 131, 97–121.
- Harvey, A. C., Ruiz, E., and Shephard, N. (1994), “Multivariate Stochastic Variance Models,” *Review of Economic Studies*, 61, 247–264.
- Hayashi, T. and Yoshida, N. (2005), “On Covariance Estimation of non-synchronously Observed Diffusion Processes,” *Bernoulli*, 11, 359–379.
- Heston, S. L. (1993), “A Closed-Form Solution for Options with Stochastic Volatility with Applications to Bond and Currency Options,” *The Review of Financial Studies*, 6, 327–343.
- Horn, R. A. and Johnson, C. R. (1991), *Topics in Matrix Analysis*, Cambridge: Cambridge University Press.
- Hsieh, D. A. (1991), “Chaos and Nonlinear Dynamics: Application to Financial Markets,” *Journal of Finance*, 46, 1839–1877.

- Huang, J.-Z. and Wu, L. (2004), “Specification Analysis of Option Pricing Models Based on Time-Changed Lévy Processes,” *Journal of Finance*, 59, 1405–1439.
- Huang, X. and Tauchen, G. (2005), “The Relative Contribution of Jumps to Total Price Variance,” *Journal of Financial Econometrics*, 3, 456–499.
- Hubalek, F. and Nicolato, E. (2005), “On Multivariate Extensions of Lévy driven Ornstein-Uhlenbeck Type Stochastic Volatility Models and Multi-Asset Options,” In preparation.
- Hull, J. and White, A. (1987), “The Pricing of Options on Assets with Stochastic Volatilities,” *Journal of Finance*, 42, 281–300.
- Jacod, J. and Shiryaev, A. N. (2003), *Limit Theorems for Stochastic Processes*, vol. 288 of *Die Grundlehren der mathematischen Wissenschaften*, Berlin Heidelberg New York: Springer.
- Jones, C. S. (2003), “The Dynamics of Stochastic Volatility: Evidence from Underlying and Options Markets,” *Journal of Econometrics*, 116, 181–224.
- Koopman, S. J., Jungbacker, B., and Hol, E. (2005), “Forecasting Daily Variability of the S&P 100 Stock Index Using Historical, Realised and Implied Volatility Measurements,” *Journal of Empirical Finance*, 12, 445–475.
- Lindberg, C. (2005), “Portfolio Optimization and Statistics in Stochastic Volatility Markets,” PhD thesis, Department of Mathematical Sciences, Division of Mathematical Statistics, Chalmers University of Technology and Göteborg University, Göteborg, Sweden.
- Lundblad, C. (2004), “The Risk Return Tradeoff in the Long-Run: 1836-2003,” Working Paper, University of North Carolina at Chapel Hill.
- Maheu, J. M. and McCurdy, T. (2002), “Nonlinear Features of Realized FX Volatility,” *Review of Economics and Statistics*, 84, 668–681.
- Marquardt, T. (2004), “Fractionally Integrated ARMA Processes in Discrete and Continuous Time,” Master’s thesis, Munich University of Technology.
- Marquardt, T. and Stelzer, R. (2007), “Multivariate CARMA processes,” *Stochastic Processes and their Applications*, 117, 96–120.
- Martens, M., Dijk, D. v., and de Pooter, M. (2004), “Modeling and Forecasting S&P 500 Volatility: Long Memory, Structural Breaks and Nonlinearity,” Working Paper, Erasmus University Rotterdam.
- Martens, M. and Zein, J. (2004), “Predicting Financial Volatility: High-Frequency Time-Series Forecasts vis-à-vis Implied Volatility,” *Journal of Futures Markets*, 24, 1005–1028.

## Bibliography

- Masuda, H. (2004), “On Multidimensional Ornstein-Uhlenbeck Processes Driven by a General Lévy Process,” *Bernoulli*, 10, 97–120.
- Merton, R. C. (1973), “Theory of Rational Option Pricing,” *Bell Journal of Economics and Management Science*, 4, 141–183.
- (1976), “Option Pricing When Underlying Stock Returns are Discontinuous,” *Journal of Financial Economics*, 3, 125–144.
- (1980), “On Estimating the Expected Return on the Market : An Exploratory Investigation,” *Journal of Financial Economics*, 8, 323–361.
- Métivier, M. (1982), *Semimartingales: a Course on Stochastic Processes*, vol. 2 of *De Gruyter Studies in Mathematics*, Berlin: Walter de Gruyter.
- Métivier, M. and Pellaumail, J. (1980), *Stochastic Integration*, New York: Academic Press.
- Müller, U. A., Dacorogna, M. M., Davé, R. D., Olsen, R. B., Pictet, O. V., and von Weizsäcker, J. E. (1997), “Volatilities of Different Time Resolutions - Analyzing the Dynamics of Market Components,” *Journal of Empirical Finance*, 4, 213–239.
- Oomen, R. C. (2005), “Properties of Bias-Corrected Realized Variance Under Alternative Sampling Schemes,” *Journal of Financial Econometrics*, 3, 555–577.
- Pan, J. (2002), “The Jump-Risk Premia Implicit in Options: Evidence from an Integrated Time-Series Study,” *Journal of Financial Economics*, 63, 3–50.
- Pedersen, A. R. (1995), “A New Approach to Maximum Likelihood Estimation for Stochastic Differential Equations Based on Discrete Observations,” *Scandinavian Journal of Statistics*, 22, 55–71.
- Pigorsch, C. and Stelzer, R. (2007), “A Multivariate Extension of the Ornstein-Uhlenbeck Stochastic Volatility Model,” Working Paper, Munich University of Technology.
- Pong, S., Shackleton, M. B., Taylor, S. J., and Xu, X. (2004), “Forecasting Currency Volatility: A Comparison of Implied Volatilities and AR(FI)MA Models,” *Journal of Banking & Finance*, 28, 2541–2563.
- Poterba, J. M. and Summers, L. H. (1986), “The Persistence of Volatility and Stock Market Fluctuations,” *The American Economic Review*, 76, 1142–1151.
- Protter, P. (2004), *Stochastic Integration and Differential Equations*, vol. 21 of *Applications of Mathematics*, Springer, 2nd ed.

## Bibliography

- Scott, L. O. (1987), “Option Pricing When the Variance Changes Randomly - Theory, Estimation, and an Application,” *Journal of Financial and Quantitative Analysis*, 22, 419–438.
- Stelzer, R. (2006), “Linear Operators on the Real Symmetric Matrices whose Exponential Preserves the Inertia,” Working Paper, Munich University of Technology.
- Tauchen, G. (2005), “Stochastic Volatility in General Equilibrium,” Working Paper, Duke University.
- Taylor, S. J. and Xu, X. (1997), “The Incremental Volatility Information in One Million Foreign Exchange Quotations,” *Journal of Empirical Finance*, 4, 317–340.
- Thomakos, D. D. and Wang, T. (2003), “Realized Volatility in the Futures Markets,” *Journal of Empirical Finance*, 10, 321–353.
- Todorov, V. and Tauchen, G. (2005), “Simulation Methods for Lévy-Driven CARMA Stochastic Volatility Models,” Working Paper, Duke University.
- Voev, V. and Lunde, A. (2007), “Integrated Covariance Estimation using High-frequency Data in the Presence of Noise,” *Journal of Financial Econometrics*, 5, 68–104.
- Zhang, L., Mykland, P. A., and Aït-Sahalia, Y. (2005), “A Tale of Two Time Scales: Determining Integrated Volatility with Noisy High-Frequency Data,” *Journal of the American Statistical Association*, 100, 1394–1411.
- Zhou, B. (1996), “High-Frequency Data and Volatility in Foreign-Exchange Rates,” *Journal of Business & Economic Statistics*, 14, 45–52.



# Eidesstattliche Versicherung

Hiermit versichere ich, dass ich die vorliegende Arbeit:

## **Estimation of Continuous–Time Financial Models Using High–Frequency Data**

selbstständig und ohne Benutzung anderer als der angegebenen Hilfsmittel angefertigt habe.

Eine Anmeldung der Promotionsabsicht habe ich an keiner anderen Fakultät oder Hochschule beantragt.

München, den 12. Juli 2007

Christian Pigorsch



# CURRICULUM VITAE

## EDUCATION

2002 - present	<b>University of Munich, Department of Statistics</b> Ph.D. Student in Statistics Supervisor: Prof. Stefan Mittnik, Ph.D.
12/2004 - 02/2005, 09/2005	<b>Duke University, Department of Economics</b> Visiting Scholar
03/2005 - 04/2005	<b>Washington University in St. Louis, Department of Economics</b> Visiting Scholar
1996 - 2002	<b>University at Kiel, Department of Economics</b> Diploma in Quantitative Economics (grade: 1.5) Diploma Thesis: "Zinsprognose anhand der Zinsstruktur—Ergebnisse für Deutschland"

## ACADEMIC AND PROFESSIONAL EXPERIENCES

since 2002	Research and Teaching Assistant at the Seminar for Financial Econometrics, Department of Statistics, University of Munich, Germany
02/1999 - 04/1999	Internship at the Hamburgische Landesbank, Germany, Empirical Derivative Research Division
1998 - 2002	Student Research Assistant at the Institute of Statistics and Econometrics, University at Kiel, Germany
1998 - 1999	Student Research Assistant at the Kiel Institute for the World Economy

## ACADEMIC AWARDS

2002	Erich-Schneider Award of the University at Kiel, Department of Economics, for Best M.S. Thesis of the Year
2002	Gerhard-Fürst Award in the Diploma and Master's Dissertations Category granted by the Federal Statistical Office of Germany